CONFERENCE MANUAL

DAY ONE

SPONSORED BY:

AMERICAN BAR ASSOCIATION
STANDING COMMITTEE ON LAW & NATIONAL SECURITY

CENTER FOR NATIONAL SECURITY LAW
UNIVERSITY OF VIRGINIA SCHOOL OF LAW

CENTER ON LAW, ETHICS AND NATIONAL SECURITY
DUKE UNIVERSITY SCHOOL OF LAW

CENTER ON NATIONAL SECURITY AND THE LAW
GEORGETOWN LAW

WASHINGTON COLLEGE OF LAW
AMERICAN UNIVERSITY

NATIONAL SECURITY LAW INSTITUTE
ANTONIN SCALIA LAW SCHOOL, GEORGE MASON UNIVERSITY

GEORGE WASHINGTON UNIVERSITY LAW SCHOOL

THE CENTER ON LAW AND SECURITY & JUST SECURITY
NEW YORK UNIVERSITY LAW SCHOOL

INSTITUTE FOR NATIONAL SECURITY AND COUNTERTERRORISM
SYRACUSE UNIVERSITY COLLEGE OF LAW

NOVEMBER 16 — NOVEMBER 17, 2017

THE CAPITAL HILTON
1001 16TH STREET, NW
WASHINGTON, D.C.
## Panel I

### National Security Law in Private Practice

<table>
<thead>
<tr>
<th>Title and Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>AML and Sanctions: 2017 Trends and Developments — Wilmer Cutler Pickering Hale and Dorr LLP</td>
<td>5</td>
</tr>
<tr>
<td>CFIUS Actions Highlight Focus on Critical Technologies and Infrastructure — Ambassador Robert Kimmitt, Benjamin Powell, Stephen Preston, and Jason Chipman</td>
<td>48</td>
</tr>
<tr>
<td>President Trump Issues Cybersecurity Executive Order — Benjamin Powell, Jonathan Cedarbaum, and Leah Schloss</td>
<td>53</td>
</tr>
<tr>
<td>Congress Seeks to Force (and Tie) President’s Hands on Sanctions Through Passage of Significant New Law Codifying and Expanding U.S. Sanctions on Russia, North Korea and Iran — Judith Lee, Adam Smith, Henry Phillips, Stephanie Connor, and Christopher Timura</td>
<td>56</td>
</tr>
<tr>
<td>National Security Practice — Gibson, Dunn &amp; Crutcher LLP</td>
<td>67</td>
</tr>
<tr>
<td>Trump Administration Prioritization of Cybersecurity — Alexander Southwell, Caroline Krass, Ryan Bergezieker, Jeana Bisnar Maute, and Casper Yen</td>
<td>78</td>
</tr>
<tr>
<td>Update on CFIUS Developments: Proposed Legislation and Reflections on CFIUS Under the Trump Administration — Covington &amp; Burling LLP</td>
<td>85</td>
</tr>
</tbody>
</table>

## Panel II

### Artificial Intelligence: Issues of Technology, Ethics, and Law

<table>
<thead>
<tr>
<th>Title and Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparing for the Future of Artificial Intelligence — Executive Office of the President National Science and Technology Council Committee on Technology</td>
<td>190</td>
</tr>
<tr>
<td>Advancing the Debate on Killer Robots: 12 Key Arguments for a Preemptive Ban on Fully Autonomous Weapons — Human Rights Watch and International Human Rights Clinic</td>
<td>264</td>
</tr>
<tr>
<td>Artificial Intelligence Policy: A Primer and Roadmap — Ryan Calo</td>
<td>291</td>
</tr>
<tr>
<td>Machines Without Principals: Liability Rules and Artificial Intelligence — David Vladeck</td>
<td>319</td>
</tr>
<tr>
<td>At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law — Lt. Col. Alan Schuller</td>
<td>344</td>
</tr>
<tr>
<td>An Introduction to Autonomy in Weapon Systems — Paul Scharre and Michael Horowitz</td>
<td>383</td>
</tr>
<tr>
<td>Great Power Competition and the AI Revolution: A Range of Risks to Military and Strategic Stability — Elsa Kania</td>
<td>406</td>
</tr>
</tbody>
</table>
### Panel III

**National Security Law and Policy, from the Privacy and Civil Liberties Oversight Board to the Department of Justice**

<table>
<thead>
<tr>
<th>Title and Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crypto Wars 2.0: Why Listening to Apple on Encryption Will Make America More Secure — Paul McLaughlin</td>
<td>410</td>
</tr>
<tr>
<td>Shedding Light on the &quot;Going Dark&quot; Problem and the Encryption Debate — John Mylan Traylor</td>
<td>441</td>
</tr>
<tr>
<td>Presidential Intelligence — Samuel Rascoff</td>
<td>468</td>
</tr>
<tr>
<td>Preventing &quot;Going Dark&quot;: A Sober Analysis and Reasonable Solution to Preserve Security in the Encryption Debate — Eric Manpearl</td>
<td>538</td>
</tr>
<tr>
<td>Regulating Domestic Intelligence Collection — Emily Berman</td>
<td>567</td>
</tr>
<tr>
<td>The Foreign Intelligence Surveillance Act of 1978</td>
<td>622</td>
</tr>
</tbody>
</table>

### Panel IV

**Cyber Operations and the Law of Armed Conflict**

<table>
<thead>
<tr>
<th>Title and Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respect for Sovereignty in Cyberspace — Michael Schmitt and Liis Vihul</td>
<td>672</td>
</tr>
<tr>
<td>Spying and Fighting in Cyberspace: What is Which? — Gary Brown</td>
<td>704</td>
</tr>
<tr>
<td>International Cybertorts: Expanding State Accountability in Cyberspace — Rebecca Crootof</td>
<td>726</td>
</tr>
<tr>
<td>Symposium on Sovereignty, Cyberspace, and Tallinn Manual 2.0: Sovereignty in the Age of Cyber — Gary Corn and Robert Taylor</td>
<td>799</td>
</tr>
<tr>
<td>Concluding Observations on Sovereignty in Cyberspace — Gary Corn and Robert Taylor</td>
<td>805</td>
</tr>
</tbody>
</table>

### Panel V

**Legal Ethics and Leaking by Government Attorneys**

<table>
<thead>
<tr>
<th>Title and Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Model Rules of Professional Conduct; Rules 1.6, 1.8, 1.9, 1.13 — American Bar Association</td>
<td>808</td>
</tr>
<tr>
<td>Financial Rewards for Whistleblowing Lawyers — Kathleen Clark and Nancy Moore</td>
<td>828</td>
</tr>
</tbody>
</table>
Panel I:

National Security Law in Private Practice

Moderator: Caroline Krass

Discussants: Robert Kimmitt, Amy Jeffress, Ingrid Price, Don Vieira
April 27, 2017

AML and Sanctions: 2017 Trends and Developments

TABLE OF CONTENTS

Executive Summary ..................................................................................................................... 1

BSA/AML Regulatory Trends and Developments ................................................................. 2
  Treasury and FinCEN.............................................................................................................. 3
  New York Department of Financial Services ........................................................................ 8
  SEC and FINRA Examination Priorities for 2017 ............................................................. 10
  Congress: The U.S. House Financial Services Committee ........................................... 10
  FATF Mutual Evaluation of the United States .................................................................... 12
  The Clearing House Advocates AML Reform .................................................................... 13

Sanctions Regulatory Trends and Developments ............................................................... 15
  Russia Sanctions .................................................................................................................. 15
  Cuba Sanctions .................................................................................................................... 17
  Iran Sanctions ...................................................................................................................... 18
  Other Significant Developments ......................................................................................... 22

Enforcement ............................................................................................................................ 25
  AML Enforcement .............................................................................................................. 25
  Sanctions Enforcement ....................................................................................................... 31
I. Executive Summary

Over the past year, regulators continued to actively examine compliance, introduce new and heightened requirements, recalibrate global priorities, and aggressively pursue enforcement. This report reviews recent trends and developments impacting financial institutions with respect to the U.S. Bank Secrecy Act/Anti-Money Laundering (BSA/AML) and economic sanctions regulatory landscape.

It remains to be seen how the arrival of the Trump Administration will impact the government’s approach and priorities in these areas. All current indications, however, are that BSA/AML and sanctions regulation are likely to remain a priority. Secretary of the U.S. Department of the Treasury (Treasury) Steven Mnuchin, for example, recently identified combating financial crime and terrorism financing as “core missions” of the Treasury in a statement supporting President Trump’s discretionary budget proposal. That proposal spares the Treasury’s financial crime units from reduced funding. In addition, other regulators at all levels continue to signal a sustained commitment to strengthening and policing the financial system’s protections against financial crime.

Regulators in 2016 focused on strengthening anti-money laundering protections. A new BSA/AML rule intended to prevent criminals, kleptocrats, and others looking to hide ill-gotten proceeds from anonymously accessing the U.S. financial system was announced by the Financial Crimes Enforcement Network (FinCEN). The rule creates new requirements for customer due diligence and identification of beneficial owners—priorities that are likely to be at the top of the examination agenda, as compliance with the new rule becomes mandatory in 2018.

Other federal and state regulators also continue to play active roles. Prudential regulators focused on BSA/AML compliance in 2016, bringing over a half-dozen enforcement actions when significant deficiencies were identified. Securities regulators also focused on BSA/AML compliance in 2016, and have already signaled that BSA/AML issues will remain a priority for 2017. And at the state level, New York played a very active role in 2016 and appears poised to continue to do so with the rollout of new rules mandating certification of compliance with AML transaction monitoring and filtering program requirements, as well as new cybersecurity regulations.

Sanctions continue to play a central role in the U.S. government’s response to geopolitical events, most prominently Iran’s nuclear program, a deteriorated security situation in eastern Ukraine, and “normalized” relations between the United States and Cuba. Already in 2017, the Trump Administration has cautioned Iran that it is “on notice” concerning ballistic missile testing, and followed that warning with new designations of individuals and entities related to those activities. While sanctions policy is always responsive to world events that can be difficult to predict, 2017 is likely to be an especially challenging year for companies seeking to identify the most relevant risks, requirements and trends likely to arise in this area of law.

As to enforcement, 2017 began with the resolution of significant enforcement actions involving both BSA/AML and sanctions violations that carried hefty penalties, included criminal charges, and were the result of parallel investigations by multiple regulators that have become the hallmark of enforcement in recent years. While the frequency and size of enforcement actions in 2016 were less than in recent years, the start of 2017 suggests the broader trend remains in the direction of continued intense scrutiny.

This report highlights the most notable BSA/AML and sanctions developments in 2016 and into the first quarter of 2017.
II. BSA/AML Regulatory Trends and Developments

The BSA/AML regulatory environment continued to evolve over the past year. Regulators addressed perceived weaknesses in the U.S. regulatory regime and other hot button issues, including a lack of transparency into the identities of beneficial owners behind accounts held by legal entities, the emerging threat posed by cybercriminals, and the de-risking of foreign correspondent accounts.

While the new Administration has signaled its intention to reduce the burden posed by financial regulation, there is little indication to date of significant plans to change course with respect to BSA/AML regulation. It therefore seems likely that regulators in the near term will continue to address transparency issues and cybersecurity, as well as other new threats that might emerge over the course of 2017. They may also respond to a number of reports and recommendations issued by lawmakers and industry groups in 2016.

To understand the issues that regulators are most likely to emphasize in the coming months, it is useful to review recent developments from last year. We begin with the most significant regulatory development of last year, which came from FinCEN—the Treasury’s lead agency for combating money laundering and safeguarding the financial system from illicit use. In the first half of 2016, and in the wake of the “Panama Papers,” FinCEN released much-anticipated new rules for customer due diligence and the identification of beneficial ownership—two priorities that are likely to remain at the top of the regulatory agenda for 2017.

Other notable developments at FinCEN include the issuance of geographic targeting orders focusing on identifying individuals behind high-end real estate; new guidance on cybersecurity; and a proposal to extend Customer Identification Program and AML requirements to banks lacking a federal functional regulator. Since Jennifer Shasky Calvery’s departure from FinCEN’s top post in May 2016, the agency has been operating under the interim leadership of Acting Director Jamal El-Hindi, who previously served as deputy director and head of FinCEN’s Policy Division.

At the state level, New York continues to play an active role. New York State’s Department of Financial Services (NYDFS), under the leadership of Maria T. Vullo, who took the helm from Benjamin Lawsky in 2016, appears committed to carrying forward New York’s aggressive regulatory and enforcement posture. NYDFS issued new rules requiring certification of compliance with AML transaction monitoring and filtering program requirements, consistent with the desire to hold individuals accountable for compliance failings. Also notable—given the increasing convergence between AML and cybersecurity issues—is NYDFS’s promulgation in February 2017 of new cybersecurity regulations requiring regulated financial institutions to establish and maintain cybersecurity programs.

Over the past year, we have seen a dramatic convergence of AML and cybersecurity and expect this trend to continue in 2017. Traditionally, financial institutions have approached cybersecurity and AML compliance separately, with different personnel and reporting lines. However, U.S. regulators now expect that financial institutions take a holistic approach to cyberthreats and incorporate such information into Suspicious Activity Reports (SARs) filed pursuant to the institution’s BSA obligations. Notably, in October 2016, the federal banking regulators issued a joint advance notice of proposed rule-making on cybersecurity regulations and efforts to improve the safety and soundness of the U.S. financial system.4

Given the increased regulatory scrutiny at the state and federal levels, financial institutions should ensure that their cybersecurity and AML compliance personnel understand when to escalate a cyber-event to AML compliance and the information needed to satisfy the relevant reporting requirements.

In 2016, securities regulators also continued to focus on AML compliance. Both the Securities and Exchange Commission (SEC) and the Financial Industry Regulatory Authority (FINRA) already have announced an intention to make AML a priority in 2017. Last year was also characterized by a number of significant policy developments that are reviewed in more detail below: Congress devoted particular attention to the threat that terrorist financing poses to the U.S. financial system; the Financial Action Task
Force (FATF) released its “mutual evaluation” assessing the strength of the U.S. regime to combat money laundering and terrorist financing; and The Clearing House called for a paradigm shift to redesign the U.S. approach to AML to better protect national security and aid law enforcement.

The following is a discussion of the key BSA/AML regulatory developments in 2016 and early 2017.

A. Treasury and FinCEN

1. New Requirements for Customer Due Diligence and Identification of Beneficial Owners

FinCEN, on May 11, 2016, released its long-awaited Final Customer Due Diligence Rule (CDD Rule) that will require certain financial institutions to “look through” nominal legal entity account holders to identify the account’s beneficial owners who own or control the entity. The rule also explicitly establishes a “fifth pillar” in FinCEN’s AML program requirement mandating that certain institutions implement risk-based procedures for conducting customer due diligence (CDD) on all customers. Compliance with the CDD Rule becomes mandatory on May 11, 2018.

The CDD Rule is intended to prevent criminals, kleptocrats and others looking to hide ill-gotten proceeds from accessing the financial system anonymously. FinCEN’s announcement of the final rule last year was the culmination of an extensive, six-year rule-making process and a key step in the U.S. government’s ongoing efforts to combat money laundering, terrorist financing and tax evasion on the heels of the Panama Papers. Publication of the final rule coincided with calls for Congress to adopt legislation that would require the collection of beneficial ownership information at the time legal entities are formed in the United States.

The U.S. Department of Justice (DOJ) recently underscored its support for the new rule, explaining that the Criminal Division “remain[s] sharply focused on understanding the ownership structure and the apparent ease with which criminal organizations and individuals use shell companies to move and ultimately conceal criminal proceeds.” The acting head of the Criminal Division recently described the rule “as a critical step toward greater transparency” and emphasized DOJ “will be taking a hard look at compliance with this rule in the course of our future investigations.”

The CDD Rule addresses a weakness in the U.S. AML regime identified by the FATF in its “mutual evaluations” of the United States—namely, shortcomings with respect to the identification and verification of the individuals associated with legal entity customers. The rule does so by requiring the identification of the “beneficial owners” of legal entity customers. Beneficial owners are defined as each individual who owns 25 percent or more of the entity, and a single individual who has significant responsibility for controlling the entity.

In addition to its beneficial ownership requirement, the CDD Rule includes CDD standards for covered financial institutions subject to AML program requirements. The so-called “fifth pillar,” which, according to FinCEN, merely formalizes existing practice, will require covered financial institutions to establish risk-based procedures to understand the “nature and purpose of the customer relationship,” and to conduct ongoing monitoring to identify and report suspicious transactions and update customer information. FinCEN emphasized that the fifth pillar does not require a continuous or periodic refresh of customer information. Rather, an institution must update customer information, including beneficial ownership, if during its normal monitoring it detects information relevant to assessing or reevaluating customer risk.

The CDD Rule applies only to “covered financial institutions”: banks, broker-dealers, mutual funds, futures commission merchants, and commodities introducing brokers, which are already subject to Customer Identification Program (CIP) requirements. FinCEN has also emphasized that the CDD Rule’s provisions are a “floor, not a ceiling,” suggesting that it may be appropriate for institutions to do more than the minimum required by the CDD Rule in circumstances of heightened risk.
Key Takeaways

➤ **Updates to Existing CDD Programs.** Many covered financial institutions already collect some beneficial ownership information and have updated their AML policies and procedures in anticipation of the CDD Rule. However, all covered financial institutions should revisit their policies, procedures and training materials to ensure their current practices meet the requirements of the CDD Rule by the May 2018 compliance deadline. While FinCEN asserts that the “fifth pillar” provisions are not new requirements, covered financial institutions may find that their procedures do not actually incorporate these expectations.

➤ **CDD Expectations for Non-Covered Financial Institutions.** Non-covered financial institutions with SAR responsibilities may want to consider establishing some form of risk-based customer due diligence processes. FinCEN and other federal functional regulators have emphasized that customer due diligence is a key input in SAR monitoring and analysis. And FinCEN's stated position that the “fifth pillar” merely formalizes existing expectations suggests that these expectations may be broadly applicable across FinCEN-regulated entities and not just applicable to covered financial institutions.

➤ **Trigger-Based Updates.** Also, while FinCEN expects financial institutions to conduct a "monitoring-triggered" update of customer information, it did not specify which triggers should be used. For example, FinCEN or prudential regulators may expect triggers to capture a change in ownership of a legal entity customer. Covered financial institutions may find themselves subject to criticism for a "silo effect" if salient information for purposes of ongoing monitoring is not effectively communicated from all relevant aspects of the firm to the AML compliance function. (This has been a frequent regulatory criticism with respect to institutions’ SAR programs.)

➤ **Financial Intermediaries.** SAR monitoring should not necessarily stop at the level of the legal entity customer, notwithstanding the fact that the CDD Rule states that intermediaries may be treated as legal entity customers under certain circumstances. In fact, failure to monitor or report underlying customer activity in intermediated accounts can attract regulatory scrutiny and lead to enforcement actions.9

➤ **Leveraging CDD for Other Compliance Efforts.** One of the stated benefits of the CDD Rule is that it may serve to enhance financial institutions’ compliance efforts in the areas of Office of Foreign Assets Control (OFAC) sanctions, currency transaction reporting requirements, tax and others. Covered financial institutions should assess the information flow among their new or updated CDD controls and the groups responsible for these other compliance requirements.

For additional details, refer to our Client Alert: FinCEN Finalizes Beneficial Ownership and Customer Due Diligence Requirements.

2. **FinCEN Focuses on High-End Real Estate**

In 2016, FinCEN turned its attention to the money laundering risks posed by high-end real estate. On January 13, 2016, FinCEN issued two Geographic Targeting Orders (GTOs), temporarily requiring certain title insurance companies to report the identities of the natural persons behind companies paying cash for high-end residential real estate in Manhattan and Miami-Dade County.10 The GTOs were applicable to residential properties with a sale price of over $1 million in Miami-Dade and over $3 million in Manhattan. They were motivated by FinCEN’s concern that individuals purchasing high-end residential real estate through limited liability companies or other legal entities and without bank financing might be trying to disguise their identity and hide assets.

The issuance of the GTOs marked the first time that the federal government required real estate companies to disclose the names of individuals behind cash transactions. Information gathered pursuant to the GTOs will be stored in a FinCEN database intended to help law enforcement identify the natural persons involved in transactions vulnerable to money laundering abuse.
The GTOs are part and parcel of a recent increase in law enforcement scrutiny of the real estate industry. The issuance of the GTOs brings the United States more in line with the United Kingdom, where real estate agents are required to submit SARs to the government when they suspect a transaction involves funds from a criminal source.

In August 2016, FinCEN extended the GTOs to six major U.S. geographic areas: (1) all boroughs of New York City; (2) Miami-Dade County, Florida, and the two counties immediately north (Broward and Palm Beach); (3) Los Angeles County, California; (4) three counties comprising part of the San Francisco area (San Francisco, San Mateo and Santa Clara counties); (5) San Diego County, California; and (6) the county that includes San Antonio, Texas (Bexar County).\(^1\)

In February 2017, FinCEN renewed the GTOs for a period extending 180 days from February 24, 2017. In its press release announcing the renewal of the GTOs, FinCEN noted: “About 30 percent of the transactions covered by the GTOs involve a beneficial owner or purchaser representative that is also the subject of a previous SAR. This corroborates FinCEN’s concerns about the use of shell companies to buy luxury real estate in ‘all-cash’ transactions.”\(^2\)

FinCEN may further increase regulatory or criminal enforcement actions in the real estate field in 2017. Depending on the results of these GTOs, it is also possible that FinCEN and other regulators may further expand AML compliance requirements in the real estate sector.

3. FinCEN Guidance on E-Mail Compromise Fraud Schemes and Cyber-Events

a) FinCEN Issues Guidance on E-Mail Compromise Fraud

On September 6, 2016, FinCEN issued an advisory to help financial institutions guard against e-mail fraud schemes wherein criminals misappropriate funds by deceiving financial institutions and their customers into conducting wire transfers.\(^3\) FinCEN warned against two main types of e-mail compromise frauds:

- Business E-mail Compromise (BEC), which targets a financial institution’s commercial customers; and
- E-mail Account Compromise (EAC), which targets a victim’s personal accounts.

BEC and EAC schemes are part of a growing trend of cyber-enabled crime affecting financial institutions. Since 2013, there have been approximately 22,000 reported cases of BEC and EAC fraud involving $3.1 billion.

FinCEN notes that both BEC and EAC schemes typically have three stages:

- First, criminals unlawfully access a victim’s e-mail account through social engineering (e.g., by tricking the victim into revealing information) or computer intrusion techniques. Criminals then exploit the victim’s e-mail account to obtain information about the victim’s financial institutions and account details.
- Second, criminals use the victim’s stolen information to e-mail fraudulent wire transfer instructions to the financial institution in a manner appearing to be from the victim. Either the criminal will use the victim’s actual e-mail account or will create a fake e-mail account that resembles the victim’s actual one, inserting an underscore between the first and last names instead of a dot, for example.
Third, the criminal will trick the victim’s financial institution into conducting unauthorized wire transfers that appear legitimate. The fraudulent transaction instructions direct the wire transfer to the criminal’s domestic or foreign bank accounts.14

Key Takeaways

 Companies should be on alert for the red flags15 indicative of BEC and EAC fraud including:

  o A customer’s seemingly legitimate e-mailed transaction instructions contain different language, timing, and amounts than previously verified and authentic transaction instructions.

  o Transaction instructions originate from an e-mail account closely resembling a known customer’s e-mail account; however, the e-mail address has been slightly altered by adding, changing, or deleting one or more characters.

  o E-mailed transaction instructions direct payment to a known beneficiary; however, the beneficiary’s account information is different from what was previously used.

 Financial institutions are encouraged to report unauthorized wire transfers to the FBI or U.S. Secret Service within 24 hours to increase the chances of recovering stolen funds.

 In addition to notifying these agencies, a financial institution may be required to file a SAR if it knows, suspects, or has reason to suspect a transaction conducted or attempted by, at, or through the financial institution involves funds derived from illegal activity, like a BEC. When filing a SAR, financial institutions should provide all pertinent available information, including cyber-related information in the SAR form and narrative.

b) FinCEN Guidance on Cyber-Events and Cyber-Enabled Crimes

On October 25, 2016, FinCEN issued an advisory to financial institutions on when to file SARs on cyber-events and cyber-enabled crimes.16 FinCEN noted that financial institutions should include available IP addresses and accompanying timestamps associated with fraudulent wire transfers being reported, even if a cyber-event was not involved in the suspicious activity. And when suspicious transactions do involve cyber-events, FinCEN advised financial institutions to include in SARs all relevant and available information regarding the suspicious transactions and the cyber-event, including the type, magnitude, and methodology of the cyber-event, and the signatures that suggest a cyber intrusion.

Financial institutions are not required to file SARs reporting the continuous scanning or probing of their systems or networks, but should file a SAR where an otherwise reportable cyber-event has been unsuccessful. Financial institutions should encourage collaboration between their AML compliance personnel and cybersecurity personnel to efficiently detect and report cyber-events.

4. In Response to De-Risking, Regulators Issue Joint Fact Sheet on Foreign Correspondent Banking

De-risking continues to be a hot topic of conversation at gatherings of AML professionals and regulators. De-risking occurs when financial institutions withdraw from certain business lines or countries that they deem too risky from an AML or sanctions perspective and can have, among other consequences, the ultimate impact of driving transactions out of the well-regulated financial systems and away from the watchful eyes of regulators and law enforcement. De-risking is of particular concern to those engaged in foreign correspondent banking. In light of high compliance costs and regulatory scrutiny, banks have become reluctant to maintain correspondent accounts for foreign financial institutions (FFIs) from certain high-risk countries.
On August 30, 2016, the Treasury and the federal banking agencies\textsuperscript{17} issued a \textit{Joint Fact Sheet on Foreign Correspondent Banking} (and corresponding Blog post) in an effort to dispel two myths about U.S. supervisory expectations that presumably drive de-risking. The Joint Fact Sheet clarifies: (1) that there is no general expectation that U.S. banks conduct due diligence on the customers of their FFI customers (i.e., no requirement to know their customers’ customers); and (2) that the AML and OFAC enforcement regime is not one of “zero tolerance.”\textsuperscript{18}

Under Section 312 of the USA PATRIOT Act and its implementing regulation, U.S. depository institutions are required to assess the money laundering risk presented by their foreign correspondent accounts. Such assessments must take into account: (1) the nature of the FFI’s business and the markets it serves; (2) the type, purpose, and anticipated activity of the account; (3) the nature and duration of the account relationship; (4) the supervisory regime of the jurisdiction in which the FFI is licensed; and (5) information about the FFI’s AML record.\textsuperscript{19} Although the Joint Fact Sheet states that there is currently no requirement for U.S. depository institutions to conduct due diligence on an FFI’s customers, banks should consider whether the due diligence information provided by their FFI customers is sufficient to fully assess the AML and sanctions risks posed by the foreign correspondent banking relationship. U.S. depository institutions may have to request additional information about the underlying activity in an FFI’s account in order to satisfy their risk-based obligations under Section 312.

Financial institutions should also note that the Joint Fact Sheet reflects the views only of the Treasury and the federal banking agencies and not that of DOJ, state and local law enforcement; state banking regulators; or securities regulators, all of whom play an important role in the AML and OFAC enforcement regime.

5. \textbf{Proposal to Extend CIP and AML Program Requirements to Banks Lacking a Federal Functional Regulator}

On August 25, 2016, FinCEN proposed a rule that would conform the AML program obligations of banks not regulated by a federal functional regulator to those of banks that are regulated by a federal functional regulator.\textsuperscript{20} The proposed rule would require banks lacking a federal functional regulator to establish and maintain an AML program that includes, at a minimum, the same five pillars as a traditional AML program: (1) internal controls; (2) independent testing; (3) a designated compliance officer; (4) training for appropriate personnel; and (5) risk-based procedures for conducting appropriate ongoing customer due diligence. In addition, the proposed rule would extend to these banks requirements to establish and implement a Customer Identification Program and to identify the beneficial owners of legal entity customers.

Banks without a federal functional regulator include state-chartered non-depository trust companies, non-federally insured credit unions, private banks, non-federally insured state banks and savings associations, and international banking entities. If finalized as proposed, the rule would fill a gap in the AML regime and subject banks, regardless of whether they have a federal functional regulator, to the full panoply of AML requirements.

Because banks without a federal functional regulator are already covered by other Bank Secrecy Act recordkeeping and reporting obligations, FinCEN anticipates that they will be able to leverage existing policies, procedures and internal controls to comply with the proposed obligations. If the rule is finalized, these banks should revise their existing AML policies, procedures, and internal controls to conform with the new minimum standards.

6. \textbf{Leadership Changes}

Treasury and FinCEN also saw recent leadership changes. On February 13, 2017, the Senate confirmed Steven Mnuchin as Secretary of the Treasury. In his \textit{written responses} to questions posed by the Senate Finance Committee, Mnuchin, a former Goldman Sachs banker and Hollywood film financier, acknowledged the "serious challenges" law enforcement faces when it is unable to determine the beneficial ownership of companies that utilize the U.S. financial system. He pledged to work with
Congress and the various entities impacted by FinCEN’s new customer due diligence requirements (discussed above) to address these challenges.21

In May 2016, Jennifer Shasky Calvery stepped down as director of FinCEN. A former prosecutor, Shasky served as director since September 2012. During that time, FinCEN overhauled its enforcement division and focused the agency’s enforcement authority on certain non-bank financial institutions including casinos, money transmitters and fintech companies. Since Shasky’s departure, FinCEN has been led by Acting Director Jamal El-Hindi, who previously led FinCEN’s policy division and served as deputy director. A new director has not yet been appointed.

B. New York Department of Financial Services

In light of the Trump Administration’s stated intention to roll back federal regulations,22 we see the potential for increased activity by state regulators and attorneys general as they step in to fill perceived gaps in regulation and enforcement. New York, in particular, appears poised to continue to play an active role with respect to institutions that fall within its jurisdiction.

The head of NYDFS, Superintendent Maria Vullo, recently expressed her view that states are “well-positioned” to respond to any uncertainty at the federal level and emphasized that “whatever happens, we’re going to continue to do our job in New York.”23 Superintendent Vullo, who was appointed last year by New York Governor Andrew Cuomo, has also indicated that New York supervisors will continue to aggressively review firms for compliance with BSA/AML obligations, cybersecurity rules and other mandates within the purview of NYDFS.24

Looking back at 2016, NYDFS appears well-positioned to do so. It continued to play an active role in enforcement, assessing several of the largest AML penalties of the year. On the regulatory front, NYDFS announced new rules requiring certification of compliance with state AML laws, which went into effect at the start of this year. And in February of this year, NYDFS finalized a first-in-the-nation cybersecurity regulation for financial institutions. These developments are described in more detail below.

1. NYDFS Issues Rule Requiring Certification of Compliance With AML Transaction Monitoring and Filtering Program Requirements

New York has taken the initiative in codifying regulatory expectations for transaction monitoring and filtering programs.

On June 30, 2016, NYDFS finalized a regulation requiring certification of compliance with AML transaction monitoring and filtering program requirements.25 The regulation imposes three main requirements on New York-regulated institutions: (1) implementation of an AML transaction monitoring program; (2) implementation of a watch-list filtering program; and (3) a certification requirement. The certification requirement reflects a continuing trend by NYDFS and federal regulators toward holding executives accountable for an institution’s perceived AML and sanctions program failures. The regulation took effect January 1, 2017, and the initial certification is due to NYDFS on April 15, 2018.

The regulation creates specific state-level obligations for AML and sanctions compliance, areas that have traditionally been left to federal oversight. In some respects, the new rule goes beyond federal law, which has never codified the specific requirements of an AML transaction monitoring or sanctions filtering program in such detail.

On its face, the regulation applies only to the following entities that are subject to NYDFS AML regulations:

- all banks, trust companies, private bankers, savings banks, and savings and loan associations chartered under the New York Banking Law;
• all branches and agencies of foreign banking corporations licensed under the New York Banking
  Law to conduct banking operations in New York; and

• all check cashers and money transmitters licensed under New York Banking Law.

However, the practical effect of the rule may be even broader. For large banking organizations that
include an institution regulated by NYDFS (such as international banks with a New York branch) and that
have enterprise-wide AML monitoring and sanctions filtering systems, the system requirements in the rule
could indirectly apply to entities not directly covered by the regulation, such as broker-dealers.

The regulation requires regulated institutions to establish and maintain a transaction monitoring program
and a filtering program.

Transaction Monitoring Program. Regulated institutions are required to maintain a reasonably designed
transaction monitoring program (either manual or automated) to monitor for potential violations of the BSA
and to comply with their suspicious activity reporting obligations. A reasonably designed program should
be based on the institution’s risk assessment and include the attributes specified in the regulation, to the
extent they are applicable.

Filtering Program. Regulated institutions are required to maintain a manual or automatic filtering
program reasonably designed to interdict transactions prohibited by OFAC sanctions programs. In
response to industry comments, the rule substituted a reasonableness standard for what appeared to be
strict liability under the proposed rule. Like the transaction monitoring program, the filtering must also be
based on the institution’s risk assessment and include the attributes specified in the regulation, to the
extent they are applicable.

Certification. The rule states that a “lack of robust governance, oversight, and accountability at senior
levels” contributed to the risk that financial institutions do not detect money laundering and other criminal
activity.26 To address this issue, the rule requires regulated institutions to adopt either an annual board
resolution, signed by each director, or a senior officer “compliance finding” to certify compliance. This
requirement reflects a continuing trend by NYDFS and federal regulators toward holding executives
accountable for AML failures. The regulation does not clarify whether a single person can satisfy the
certification requirement or whether multiple signers may be required to cover all relevant areas.

Notably, the certification lacks a materiality standard, unlike the certification requirement of Sarbanes-
Oxley. The board or senior officer must certify that the systems comply with the substantive and
subjective requirements for AML transaction monitoring and filtering programs, not just that they are
reasonably designed to detect money laundering and to block sanctioned transactions.

Although the regulation does not expressly reference criminal penalties, it states that it will be enforced
pursuant to NYDFS’s “authority under any applicable laws.” NYDFS retains the authority to impose civil
monetary and equitable sanctions and to refer matters to the New York attorney general for additional
civil or criminal enforcement.27

For additional details, refer to our Client Alert: NYDFS Issues Final Rule Requiring Certification of
Compliance With AML Transaction Monitoring and Filtering Program Requirements.

2. NYDFS Issues Cybersecurity Regulations for Financial Institutions

New York has also taken a leading role in promulgating a first-in-the-nation cybersecurity regulation
applicable to banks, insurance companies, and certain other financial institutions regulated by NYDFS.28
Entities covered by the regulation will have 180 days from the effective date (March 1, 2017) to come into
compliance with most requirements, though certain provisions allow up to two years after the effective
date.29
First proposed in September 2016 and revised after two rounds of public comment, the regulation establishes requirements that in some respects duplicate federal data security obligations for financial institutions, but in some important respects differ from and go beyond federal requirements. Notably, the regulation defines "Nonpublic Information" more broadly than the definition of "customer information" under the federal Interagency Guidelines Establishing Information Security Standards.30

The regulation imposes (1) obligations to report cybersecurity incidents to NYDFS; (2) an annual certification requirement concerning compliance with the regulation; (3) requirements concerning oversight of third-party service providers; (4) obligations concerning use of multifactor authentication and encryption; and (5) requirements concerning audit trail maintenance and document destruction.

For additional details, refer to our Client Alert: New York Finalizes Cybersecurity Regulations for Financial Institutions.

C. SEC and FINRA Examination Priorities for 2017

The SEC and FINRA have announced their examination priorities for 2017, and AML continues to be a focus for both securities industry regulators.31

The SEC’s priorities for 2017 are similar to last year’s priorities. They include assessing:

- whether broker-dealers’ AML programs are tailored to specific risks;
- how broker dealers monitor for suspicious activity;
- the effectiveness of independent testing; and
- compliance with SAR requirements, including the timeliness and completeness of SARs.32

 Whereas the SEC specifically cited AML examinations of clearing and introducing brokers in its 2016 priorities letter, there was no mention of clearing or introducing brokers in the AML section of the 2017 priorities letter.

FINRA will also continue to focus on AML in 2017, especially on areas where they “have observed shortcomings,” including gaps in surveillance systems caused by data integrity problems, poorly set parameters and surveillance patterns that do not capture potentially suspicious activity.33 In particular, FINRA will continue to focus on microcap activity. It will also focus on foreign currency transactions and transactions that flow through suspense accounts, as well as controls around accounts held by nominee companies. FINRA noted that firms may perform AML monitoring using the same trade surveillance systems they use for supervisory purposes; but if they do, those systems must also monitor for the firm’s AML red flags.

D. Congress: The U.S. House Financial Services Committee

Congress is also playing an active role in exploring AML issues and proposing potential new measures. In 2016, the 114th Congress held hearings and introduced several bills pertaining to AML that are likely to serve as a starting place for the 115th Congress to revisit existing policies, assess effectiveness, and propose potential changes.

On December 20, 2016, the U.S. House Financial Services Committee’s Task Force to Investigate Terrorism Financing released a bipartisan report, Stopping Terror Finance: Securing the U.S. Financial Sector, in connection with the culmination of its two-year investigation into the threats posed by terror financing to the U.S. financial system.34 The Task Force, composed of 21 committee members, held 11 hearings from a range of expert witnesses on the terrorist financing threat and current response efforts. Topics included the nexus between terrorism, corruption and transnational crime; Iran's terrorist financing
capabilities; and terrorist financing methodologies. Based on the hearings, the Report details terrorist financing methods, key terrorist financiers such as ISIS and Boko Haram, the historical and current U.S. policy responses to terrorist financing, and methods of moving terrorist proceeds.

The Report’s long-term recommendations include improved interagency coordination and efficiency, enhanced leveraging of suspicious activity reports and information flow between government and industry, an increased number of Treasury attaches internationally, continued focus on helping developing countries combat financial crimes, and a dedication to end trade-based money laundering, among other government and policy-driven initiatives.

Prior to the Report’s publication, Task Force leaders introduced a package of five bills, none of which were passed before the end of the 114th Congress, but which may serve as a starting point for future legislation. The bills are summarized below:

- **H.R. 5594**, the “National Strategy for Combating Terrorist, Underground, and Other Illicit Financing Act,” was sponsored by Task Force Chairman Michael Fitzpatrick (R-PA) and Reps. Kyrsten Sinema (D-AZ) and Nydia Velazquez (D-NY). The bill would require the President, acting through the Treasury Secretary, to develop and publish an annual whole-of-government strategy to combat money laundering and terrorist financing. The bill passed the House on July 11, 2016, by voice vote. The Senate took no action.

- **H.R. 5602**, sponsored by Task Force Ranking Member Stephen Lynch (D-MA) and Rep. Peter King (R-NY), required more detailed information to be reported to the Treasury regarding certain types of transactions in a specific area for a limited amount of time. The bill passed the House on July 11, 2016, by a vote of 356-47, and passed the Senate December 10, 2016, with an amendment by voice vote. The House did not act on the Senate amendment prior to adjournment.

- **H.R. 5607**, the “Enhancing Treasury’s Anti-Terror Tools Act,” was sponsored by Task Force Vice Chairman Robert Pittenger (R-NC) and Ranking Member Lynch (D-MA). The bill would enhance Treasury’s anti-illicit finance tools by addressing issues that came up repeatedly in Task Force hearings. The bill passed the House on July 11, 2016, by a vote of 362-45. The Senate took no action.

- **H.R. 5603**, the “Kleptocracy Asset Recovery Rewards Act,” was sponsored by Ranking Member Lynch (D-MA) and Rep. Keith Rothfus (R-PA). The bill would establish a reward program aimed at helping the U.S. identify, freeze, and, if appropriate, repatriate assets linked to foreign government corruption, which is often an enabler of terrorism. The bill saw no action.

- **H.R. 5606**, the “Anti-Terrorism Information Sharing Is Strength Act,” was sponsored by Vice Chairman Pittenger (R-NC) and Financial Services Committee Ranking Member Maxine Waters (D-CA). The bill would refine “safe harbors” for the sharing of anti-terror information, reaffirming congressional intent in an existing statute that encourages the government to share terror methodologies with banks to help them better recognize such activity. The House tried to pass the bill under suspension of the rules (a procedure typically reserved for quick consideration of non-controversial legislation), but it failed to obtain the two-thirds vote needed to pass. The House took no further action.

Building off the work done by the Task Force, the House Financial Services Committee created a new Subcommittee on Terrorism and Illicit Finance to assist the new Administration in identifying ways to end terrorist financing. The subcommittee is chaired by Representative Stevan Pearce (R-NM). Ed Perlmutter (D-CO) is the ranking member.
E. FATF Mutual Evaluation of the United States

The Financial Action Task Force is an independent, intergovernmental body that develops and promotes policies to protect the global financial system against money laundering, terrorist financing and the financing of proliferation of weapons of mass destruction. On December 1, 2016, the FATF published its Mutual Evaluation Report of the United States. The Report presents a detailed assessment of the anti-money laundering and combating the financing of terrorism (AML/CFT) measures in place in the United States as of early 2016, when the FATF conducted its on-site visit. The FATF’s mutual evaluations influence U.S. regulatory policy by identifying deficiencies in the AML/CFT regime. The 2006 mutual evaluation, which concluded that the United States was non-compliant with FATF standards regarding beneficial ownership, spurred efforts to strengthen beneficial ownership standards, including FinCEN’s promulgation of the CDD Rule, discussed above.

The generally positive 2016 Mutual Evaluation Report recognizes the overall effectiveness of the U.S. AML/CFT regime at combating money laundering and terrorist financing, and credits the robust supervision of the banking and securities sectors while acknowledging the strong supervisory focus placed on the casino industry in recent years. The Report also identifies several weaknesses in the U.S. AML/CFT regime, including the lack of a requirement to disclose to the government the identity of beneficial owners of a company when the company is formed and there is minimal AML/CFT regulatory coverage of investment advisers, lawyers, accountants, real estate agents, and trust and company service providers (other than trust companies).

In addition to a narrative assessment of U.S. AML/CFT measures, the Report contains discrete ratings of the United States’ effectiveness and of its technical compliance with the FATF’s 40 Recommendations. The FATF gave the United States high ratings on four of the eleven effectiveness ratings (namely those relating to asset forfeiture, combating terrorist financing, and counter-proliferation sanctions). On technical compliance with the 40 Recommendations, the FATF rated the United States fully compliant with nine recommendations, largely compliant with 21 recommendations, partially compliant with six recommendations, and non-compliant with four recommendations. The scoring reflects the fact that U.S. AML/CFT laws and regulations do not always align with the corresponding FATF recommendations.

Key Takeaways

- **Beneficial Ownership.** Perhaps the most significant gap in the U.S. AML/CFT framework, according to the Report, is the lack of transparency into the beneficial ownership of legal persons. Despite the known money laundering risk posed by legal persons and arrangements, gaps in the U.S. legal framework impede law enforcement’s timely access to beneficial ownership information. The CDD Rule, discussed above, will likely address some, but not all, of the FATF’s concerns regarding beneficial ownership. The Report recommended that the United States take steps to ensure that beneficial ownership information of U.S. legal persons is timely made available to authorities by (1) requiring its collection at the federal level; and (2) requiring states to gather beneficial ownership information at the incorporation stage. Implementation of these recommendations would require legislative action by Congress and state legislatures.

- **Investment Advisers.** The Report recommended applying AML/CFT obligations to investment advisers. While some investment advisers are indirectly covered through their association with banks, bank holding companies, and/or broker dealers, investment advisers that operate outside the AML/CFT regime present a significant vulnerability. Applying AML rules to investment advisers directly would better address the vulnerability. FinCEN’s proposed investment adviser rule was pending at the time of the FATF on-site visit, and is still pending.

- **Other Under-Supervised Sectors.** The Report identified oversight gaps for non-financial businesses and professions with the notable exception of casinos and dealers in precious metals and stones, which are subject to AML requirements. The Report recommended that the United States conduct a vulnerability analysis of the non-financial businesses and professions to consider what measures could be introduced to address the higher risks to which these sectors...
are exposed. On the basis of that analysis, the Report recommended that the United States consider applying AML/CFT obligations to lawyers, accountants, and trust and company service providers (other than trust companies).

- **State-Level Monitoring.** Finding no uniform state-level AML efforts, the Report recommended that the United States improve the visibility of state-level AML activities and statistics, including through improved data-sharing. Enhanced collection of state-level money laundering investigations would allow federal authorities to better assess whether law enforcement activities at the state and local levels are consistent with federal AML/CFT priorities.

- **Section 314 Information Sharing Program.** Section 314 of the USA PATRIOT Act facilitates information exchange among financial institutions and between financial institutions and government authorities. The Report noted the importance of the Section 314 authorities and encouraged FinCEN to expand its use of Section 314(a), which allows federal, state, local, and foreign law enforcement agencies to reach out, through FinCEN, to points of contact at tens of thousands of financial institutions to locate accounts and transactions of persons that may be "engaged in or reasonably suspected . . . of engaging in terrorist acts or money laundering activities," with respect to a particular criminal investigation.\(^{38}\)

- **SAR Timing and Thresholds.** The Report found that there were technical gaps that limited the information available to competent authorities at any given time, including the application of reporting thresholds for SARs. The Report recommended a focused risk review of the existing SAR reporting thresholds and the 30/60-day window in which to report suspicious activity. The Report also recommended that FinCEN issue guidance to clarify the scope of the immediate SAR reporting requirement to make clear that the requirement to immediately notify law enforcement of certain urgent circumstances, such as terrorist activity or ongoing money laundering, applies even below the otherwise applicable thresholds. In sum, the Report may be a chance for FinCEN to revisit these thresholds and expand certain SAR programs.

F. **The Clearing House Advocates AML Reform**

On February 16, 2017, The Clearing House, a banking trade association, published a report advocating a series of reforms to the U.S. AML/CFT regime. The report, titled "A New Paradigm: Redesigning the U.S. AML/CFT Framework to Protect National Security and Aid Law Enforcement," is the product of two closed-door symposia that convened approximately 60 leading experts from government, industry and academia to identify problems with the current regime and review potential solutions.\(^{39}\)

The Clearing House notes in the report that U.S. financial firms, which are effectively deputized to identify and report money laundering and terrorist financing, spend billions of dollars each year on AML compliance, yet "many if not most of the resources devoted to AML/CFT by the financial sector have limited law enforcement or national security benefit" and may in fact damage other vital interests.\(^{40}\)

The report identifies the following reforms for immediate action:

- The Treasury, through its Office of Terrorism and Financial Intelligence (TFI), should take a more prominent role in coordinating AML/CFT policy across the government.

- FinCEN should reclaim sole supervisory responsibility for large, multinational financial institutions that present complex supervisory issues.

- TFI and FinCEN should create a robust and inclusive annual process to establish AML/CFT priorities.
• Congress should enact legislation, already pending in various forms, that requires the reporting of beneficial owner information at the time of incorporation, preventing the establishment of anonymous companies.

• TFI should strongly encourage innovation, and FinCEN should propose a safe harbor rule allowing financial institutions to innovate in a “sandbox” without fear of examiner sanction.

• Policymakers should incentivize banks to work on investigations and reporting of activity of high law enforcement or national security consequence.

• Policymakers should further facilitate the flow of raw data from financial institutions to law enforcement to assist with the modernization of the current AML/CFT technological paradigm.

• Regulatory or statutory changes should be made to the safe harbor provision in the USA PATRIOT Act (Section 314(b)) to further encourage information sharing among financial institutions, and the potential use of utilities to allow for more robust analysis of data.

• Policymakers should enhance the legal certainty regarding the use and disclosure of SARs.
III. Sanctions Regulatory Trends and Developments

Sanctions have played an increasingly central role in the U.S. government's response to Iran's nuclear program, a deteriorated security situation in eastern Ukraine, improved relations with Cuba, and other geopolitical events. The 2016 U.S. presidential election, and the new Administration, present numerous questions about how the U.S. approach to sanctions may shift, especially in light of the prominence of the Iran nuclear deal in President Trump's 2016 campaign and controversies over improving U.S.-Russia relations. Already in 2017, the Trump Administration has cautioned Iran that it is "on notice" concerning ballistic missile testing, and followed that warning with new designations of individuals and entities related to those activities. And, before its end, the Obama Administration used cyber sanctions for the first time to target Russian individuals and entities alleged to have made cyber-enabled attacks on the U.S. election system.

It is valuable, at this important juncture for U.S. sanctions policy, to reflect on the key developments in U.S. sanctions over the past year. In addition to potentially dramatic changes by the Trump Administration to Iran, Russia, and other sanctions programs, this may also be a year in which Congress takes a more active role in shaping U.S. sanctions. While sanctions policy is always responsive to world events that can be difficult to predict, 2017 is likely to be an especially challenging year for companies seeking to identify the most relevant risks, requirements and trends likely to arise in this area of law.

A. Russia Sanctions

1. Ukraine/Crimea

The United States and the European Union have maintained sanctions against Russia since 2014 in response to Russia's annexation of Crimea and its apparent support to separatist movements in eastern Ukraine. These sanctions fall into several categories.

First, pursuant to Executive Orders 13,660 and 13,661, the Treasury Department’s Office of Foreign Assets Control has added Russian and Ukrainian individuals and entities to its list of Specially Designated Nationals and Blocked Persons (SDN List). These designations began in 2014 and have continued at irregular intervals since that time, including new designations as recently as December 20, 2016.

Second, pursuant to Executive Order 13,662, OFAC has developed a novel sectoral sanctions program that targets certain types of transactions (concerning access to U.S. capital markets and access to U.S.-origin goods, technology, and services related to nonconventional energy projects) involving designated Russian firms in certain targeted industries (financial services, energy, and defense). These sectoral sanctions designations likewise began in 2014 and have continued through the end of 2016.

Third, pursuant to Executive Order 13,685, OFAC has administered an effective trade embargo with the Crimea region of Ukraine and has made a number of additions to the SDN List of entities and vessels operating there. These, too, have occurred through the end of 2016.

Fourth, the U.S. Departments of Commerce (Commerce) and State (State) have imposed heightened export controls on exports, re-exports, and transfers of certain U.S.-origin goods, software, technology and services targeting the Russian energy and defense sectors. There were no significant developments related to export controls in 2016.

OFAC also issued two general licenses under the Russia sanctions program in 2016. In September, OFAC issued General License No. 10, authorizing certain transactions with entities designated under Executive Order 13,685, the Crimea embargo, necessary to divest or transfer holdings in those entities. And in December, OFAC issued General License No. 11, authorizing certain transactions with FAU Glavgosekpertiza Rossii in connection with obtaining project design reviews or permits from the Russian federal agency with authority for approving certain infrastructure projects in Russia.
The election of President Trump has raised questions about the future of this sanctions program, as the existing multilateral commitment to maintain Russia sanctions unless and until Russia meets its commitments under the Minsk II cease-fire could be undermined if the Trump Administration pursues “a rapprochement with Putin.” The President and members of his Administration have suggested that rather than ease sanctions in exchange for Russia’s adherence to Minsk II and its cessation of activities in Ukraine, the United States may provide sanctions relief in exchange for Russian cooperation on combating global terrorist threats such as ISIS or reducing its nuclear weapons capabilities.

Yet such suggestions have been followed by reports that President Trump “may shelve, at least temporarily, his plan to pursue a deal with Moscow on the Islamic State group and other national security matters.” And the White House and members of the Trump Administration have also signaled that, in fact, the easing of sanctions against Russia would remain subject to its adherence to Minsk II and the abandonment of Crimea. Indeed, on April 21, 2017, Secretary Steve Mnuchin specifically stated that the Department of the Treasury would not authorize U.S. companies to proceed with currently prohibited energy projects in Russia. Amid this uncertainty, there is strong support for sanctions among both Democratic and Republican leaders on Capitol Hill, where any effort by the Trump Administration to ease sanctions could be met with legislative action to “backfill” existing measures currently in place under executive action. In January, a bipartisan group of senators introduced legislation that would effectively codify and escalate (including by targeting U.S. and non-U.S. financial institutions that participate in Russian sovereign debt offerings) sanctions against Russia imposed by the Obama Administration.

In December 2016, the European Council extended the application of the EU’s equivalent sectoral sanctions until July 31, 2017, and the EU is expected to renew its asset freezes and travel bans against targeted Russian individuals and entities by mid-March 2017.

2. Cyber Sanctions

In December 2016, President Obama amended a previously issued executive order designed to address cyber-enabled activities in response to “Russia’s cyber activities [that] were intended to influence the election, erode faith in U.S. democratic institutions, sow doubt about the integrity of our electoral process, and undermine confidence in the institutions of the U.S. government.” At the same time, OFAC added five entities and four individuals to the SDN List in connection with what the White House described as malicious Russian cyber activity.

These new cyber sanctions were the first of their kind, targeting Russian intelligence services and their top officers, including the Federal Security Service, or FSB, as well as three companies that the Obama Administration identified as supporting Russia’s cyber-enabled activities through intelligence activities, special training, and other capabilities.

The December 2016 action represented the first (and, thus far, only) use of the cyber sanctions, which President Obama promulgated in an April 2015 executive order in response to a number of apparent cyber-enabled threats to U.S. interests. Congress has previously authorized sanctions to combat cyber-related economic or industrial espionage, and these measures paralleled a separate executive order imposing sanctions against North Korea for “its destructive, coercive cyber-related actions during November and December 2014.”

In February 2017, OFAC announced a new general license authorizing U.S. and non-U.S. companies exporting U.S.-origin technology to the Russian Federation to engage the Russian security service in regulatory approval processes, provided that the export is otherwise licensed by the Commerce Department, the payment of any regulatory fees to the Federal Security Service does not exceed $5,000 per calendar year, and the Federal Security Service is not the end user of the exported technology.

For additional detail on the general license, refer to our Client Alert: Treasury Department Imposes New Sanctions Against Iran and Clarifies Russia Cyber Sanctions.
3. Magnitsky Act

OFAC administers sanctions pursuant to the Sergei Magnitsky Rule of Law Accountability Act of 2012 and, in 2016, added 12 individuals to the SDN List because of their involvement in human rights abuses in Russia.

B. Cuba Sanctions

2016 saw the continued normalization of U.S.-Cuba relations, a policy shift set in motion by the Obama Administration in December 2014. President Trump has sent mixed signals about whether to continue that shift. He appeared to support the concept of liberalized U.S.-Cuba trade during last year’s presidential election, but said that “we should have made a better deal.”57 The Trump Administration, like its predecessor, would require congressional cooperation to lift the embargo if it decided to significantly expand on existing sanctions relief.

But, more recently, President Trump and his representatives have emphasized the need for rejecting what he has called a “very weak agreement,” including, if necessary, by revoking executive orders signed by President Obama that gave effect to the change in policy. Upon the death of Fidel Castro, the then-President-Elect affirmed that he would “terminate [the] deal” if Cuba proved “unwilling to make a better deal for the Cuban people.”58

As in the case of Russia, the Obama Administration accomplished its sanctions objectives primarily through executive action; the Trump Administration, should it decide to reverse existing policies, could do so unilaterally by revoking executive orders.

1. January Round

In January 2016, Treasury and Commerce took action to implement President Obama’s initiatives to normalize U.S.-Cuba relations and to “engage and empower the Cuban people.” These included:59

- removing certain payment and financing restrictions for authorized exports and re-exports to Cuba of non-agricultural items or commodities;
- permitting blocked space, code-sharing, and leasing arrangements with Cuban airlines to further facilitate authorized travel to Cuba;
- authorizing additional travel-related and other transactions incident to the temporary sojourn of aircraft and vessels; and
- expanding travel authorizations related to professional meetings and other events, disaster preparedness and response projects, and information and information materials, including transactions incident to professional media or artistic productions, in Cuba.

2. March Round

In March 2016, Treasury and Commerce, in coordination with the State and Transportation Departments, announced new regulations that allowed scheduled air service between the United States and Cuba.60 The regulations also included an expansion of Cuban and Cuban nationals’ access to U.S. financial institutions and dollar-denominated transactions. In particular, OFAC allowed U.S. banks to process U-turn transactions in which Cuba or a Cuban national has an interest; to process U.S. dollar monetary instruments presented indirectly by Cuban financial institutions; and to open and maintain bank accounts in the United States for Cuban nationals in Cuba to receive payments for authorized or exempt transactions and to remit such payments back to Cuba.
3. October Round

In October 2016, Treasury and Commerce announced the sixth round of eased sanctions and export controls designed to bolster trade between the United States and Cuba. These regulatory changes, like the five preceding rounds, reflected the exercise of presidential authority to reduce existing restrictions on U.S. dealings with Cuba within the constraints of a statutory embargo that only Congress can lift. They also coincided with the release of a new Presidential Policy Directive that set forth an interagency strategy to continue with the normalization of U.S.-Cuba relations. The Trump Administration has yet to provide clarity over whether it intends to continue to pursue normalization, or whether it will instead freeze or roll back the eased sanctions that occurred under President Obama.

For additional detail about the October round of Cuba sanctions relief, refer to our Client Alert: US Implements Sixth Round of Eased Sanctions.

C. Iran Sanctions

President Trump vowed during the 2016 presidential campaign to “dismantl[e] the [disastrous] nuclear deal with Iran.” The P5+1 (China, France, Germany, Russia, the United Kingdom and the United States), the European Union and Iran had agreed to the Joint Comprehensive Plan of Action (JCPOA) in July 2015, and implemented it in January 2016.

In particular, on January 16, 2016, the P5+1 announced the arrival of “Implementation Day” for the JCPOA. In return for Iran meeting certain nuclear benchmarks, the United States and the EU implemented measures to lift the “nuclear-related” trade and financial sanctions against Iran pursuant to the JCPOA Sanctions Annex, superseding interim sanctions relief that had been in place since 2013. Although the EU has lifted nearly all sanctions against Iran, most restrictions applicable to U.S. persons and firms remain in effect.

The United States provided sanctions relief through the issuance of a new executive order, new general licenses, a new statement of licensing policy, and new guidance and FAQs; the adoption of several statutory waivers; and the delisting of various individuals and entities from the SDN List and other sanctions lists. These measures lifted nearly all secondary sanctions applicable to the activities of non-U.S. persons relating to Iran, as well as restrictions on the Iran-related activities of non-U.S. entities that are owned or controlled by U.S. persons.

But “primary” U.S. sanctions against Iran that are applicable to U.S. persons and firms, in addition to sanctions targeting Iran’s support for terrorism, human rights abuses, ballistic missile development and destabilizing regional activities, remain in effect. The prohibition on “facilitation” by U.S. persons of the Iran-related transactions of non-U.S. persons also remains in effect, albeit with limited exceptions related to the alteration of internal company policies and procedures and use of global business support systems in the context of foreign entities owned or controlled by U.S. persons, as explained more fully below.

In the EU, sanctions relief took the form of an EU Council decision implementing legislation on sanctions relief, including the delisting of Iran-related persons and entities from its sanctions lists. In contrast to U.S. sanctions relief, EU relief was relatively comprehensive by lifting most restrictions applicable to EU persons and entities, although certain sanctions relating to proliferation, missile technology, human rights and terrorism remain in effect.

These U.S. and EU measures created a variety of new compliance challenges for firms in the United States, EU and elsewhere. Most prominently, Implementation Day marked a significant divergence between U.S. and EU sanctions requirements, which now poses unique risks to firms with global operations.

Financial institutions and their risk appetites continue to play a pivotal role in the practical implementation of this sanctions relief. Indeed, since Implementation Day, the willingness of global firms to take
advantage of the new opportunities in Iran created by the sanctions relief has been mixed. The continued
limits placed on the U.S. financial system presents banking and financing challenges for non-U.S. firms
that limit their ability to operate in or with Iran. Non-sanctions risks, including those arising from the
integrity of the Iranian financial system, anti-corruption compliance and related risks, contribute to making
Iran a very challenging market in which to operate, even when U.S. or EU sanctions relief now allows it.

The election of President Trump, who has been a sharp critic of the JCPOA, raises the likelihood that
sanctions will be re-imposed and/or more aggressively enforced. In the first weeks of the Trump
Administration, OFAC imposed new non-nuclear related sanctions against Iran in response to its ballistic
missile testing. While these were limited only to new SDN designations, they signaled the arrival of rising
rather than dissipating hostilities between the United States and Iran. Indeed, in April 2017, the State
Department certified to Congress that Iran was in compliance with the JCPOA as required by the Iran
Nuclear Agreement Review Act of 2015. However, the State Department also stated at that time that the
National Security Council will lead an interagency review of the JCPOA to determine whether it is “vital to
the national security interests of the United States.” Notably, the State Department certification
concerning the JCPOA raised specific concerns over one non-nuclear related Iranian activity: Iranian
support for terrorism.

For additional detail concerning the JCPOA, summarized below, refer to our Client Alert: Iran Nuclear
Sanctions Relief Implemented.

1. U.S. Sanctions Relief

   **New Executive Order**

   On Implementation Day, the United States issued a series of statutory waivers and a new executive order
   that revoked or amended the following previous executive orders:

   - Executive Order 13,574 of May 23, 2011 (authorizing the implementation of certain sanctions
     set forth in the Iran Sanctions Act of 1996, as amended);
   - Executive Order 13,590 of November 20, 2011 (authorizing the imposition of certain sanctions
     with respect to the provision of goods, services, technology, or support for Iran’s energy and
     petrochemical sectors);
   - Executive Order 13,622 of July 30, 2012 (authorizing additional sanctions with respect to Iran);
   - Executive Order 13,645 of June 3, 2013 (authorizing the implementation of certain sanctions set
     forth in the Iran Freedom and Counter-Proliferation Act of 2012 and additional sanctions with
     respect to Iran); and
   - Executive Order 13,628 of October 9, 2012 (authorizing the implementation of certain sanctions
     set forth in the Iran Threat Reduction and Syria Human Rights Act of 2012 and additional
     sanctions with respect to Iran).

   **General Licensing for Foreign Entities Owned or Controlled by U.S. Persons**

   OFAC issued General License H (GL H) on Implementation Day, authorizing foreign subsidiaries and joint
   ventures of U.S. firms to engage in Iran-related business. GL H is not limited to specific sectors, but
   OFAC’s Guidance and FAQs accompanying the publication of GL H emphasize that authorized activities
do not include those involving SDNs; U.S.-origin items controlled for export; or any Iranian Government
   military, intelligence or law enforcement entity, among other restrictions.

   Notably, GL H does not authorize U.S. persons to directly engage in any Iran-related activities, and the
general prohibition on facilitation by U.S. persons of the Iran-related activities of non-U.S. persons—such
as by approving, financing or guaranteeing such activities—remains in effect. However, GL H provides for certain limited exceptions from the prohibition on facilitation by authorizing U.S. persons to engage in an initial set of activities to give effect to GL H:

- activities related to the establishment or alteration of operating policies and procedures of a U.S. entity or a U.S.-owned or -controlled foreign entity, to the extent necessary to allow a U.S.-owned or -controlled foreign entity to engage in transactions authorized by GL H; and

- activities to make available to those foreign entities that the U.S. person owns or controls any automated and globally integrated computer; accounting; e-mail; telecommunications; or other business support system, platform, database, application or server necessary to store, collect, transmit, generate or otherwise process documents or information related to authorized transactions.

OFAC has clarified that these exceptions to the prohibition on facilitation are intended to authorize the involvement of U.S. persons who are “board members, senior management, and employees of either a U.S. parent company or a U.S.-owned or -controlled foreign entity in the establishment or alteration of operating policies and procedures,” as well as “outside legal counsel or consultants to draft, alter, advise, or consult on such operating policies and procedures.” It also authorizes the provision by U.S. persons of “training, advice, and counseling on the new or revised operating policies and procedures.”

GL H does not authorize U.S. persons to become involved in the ongoing Iran-related operations or decision making of U.S.-owned or -controlled foreign entities, including their day-to-day operations. GL H also does not authorize U.S. persons to facilitate any activity by a foreign entity that is not U.S.-owned or -controlled, including even initial alterations to those entities’ policies and procedures.

On January 12, 2017, OFAC published “Guidance on the Provision of Certain Services Relating to the Requirements of U.S. Sanctions Laws,” which clarified that U.S. persons may provide information and guidance regarding the requirements of U.S. sanctions laws, and may opine on the legality of specific transactions, regardless of whether the U.S. person himself or herself could engage in those transactions. While the guidance is not specific to GL H or Iran sanctions, it does help to delineate the role that U.S. compliance and legal service providers can play with respect to activities by non-U.S. companies.

General License for the Importation of Iranian Foodstuffs and Carpets

OFAC issued a new general license on Implementation Day authorizing imports into the United States of, and dealings in, Iranian-origin carpets and certain foodstuffs, including pistachios and caviar.

Statement of Licensing Policy Relating to Iranian Commercial Passenger Aircraft

OFAC also issued a new Statement of Licensing Policy, pursuant to which U.S. or non-U.S. firms may request specific authorization to engage in transactions for the sale or export of commercial passenger aircraft, spare parts and components, and associated services to Iran. Unlike previous licensing policies in this sector, the new policy does not require that the transaction be focused solely on the safety of civil aviation in Iran but allows for commercial sales more broadly. It was followed, in March 2016, by new clarifying FAQs and the issuance of General License I, which authorized certain transactions related to the negotiation of, and entry into, contingent contracts for activities eligible for authorization under the Statement of License Policy.

In July 2016, OFAC issued General License J, which authorized the re-exportation of certain civil aircraft to Iran on temporary sojourn and related transactions. In December 2016, OFAC amended that general license to allow the temporary re-export of eligible civil aircraft to Iran involving code sharing arrangements. These followed new FAQs in April 2016 in which OFAC addressed the payments or the facilitation of payments to Iranian civil aviation authorities for overflights of Iran or landing in Iran.
Delisting from the SDN and Other Sanctions Lists

On Implementation Day, OFAC removed more than 400 Iranian and non-Iranian individuals and entities from the SDN List, Foreign Sanctions Evaders List, and Non-SDN Iran Sanctions Act List, as set forth in Annex II of the JCPOA.74

Many individuals and entities remain on the SDN List and therefore subject to asset freezes under various sanctions authorities. The Government of Iran and certain Iranian financial institutions remain SDNs pursuant to Executive Order 13,599 and other legal authorities, and U.S. persons must continue to block the property and interest in property of all such entities unless otherwise authorized. Secondary sanctions, by contrast, no longer apply to such Executive Order 13,599 designees, and OFAC published a new Executive Order 13,599 List that assists with compliance by identifying entities that qualify as the “Government of Iran” or an “Iranian financial institution.”75

2. EU Sanctions Relief

The EU adopted Council Decision (CFSP) 2016/37 on Implementation Day to give effect to the October 2015 Council Decision (CFSP) 2015/1863, setting forth the terms of the EU’s Iran sanctions relief.76 The European Council also published a comprehensive Information Note on EU sanctions relief.

In contrast to U.S. sanctions relief, EU relief is relatively comprehensive, lifting sanctions on nearly all financial and commercial dealings between the EU and Iran. These include financial transfers to and from Iran, including use of the SWIFT financial messaging services by non-listed Iranian entities. Certain non-listed Iranian banks are now able to open branches, subsidiaries or representative offices in the EU, and EU financial institutions may also open offices in Iran. The EU has also lifted nearly all sanctions on the Iranian oil, gas and petrochemical sectors; the shipping and shipbuilding sectors; the precious metals sector; and others.

Like the United States, the EU removed many Iranian and non-Iranian individuals and entities from its sanctions list, as set forth in Annex II of the JCPOA, and continues to apply certain limited types of sanctions against Iran, including an arms embargo and sanctions related to missile technology proliferation, human rights abuses, anti-terrorism and Iran’s destabilizing regional policies.

3. Continuing Sanctions Against Iran

Despite the broad easing of nuclear-related sanctions against Iran, a variety of restrictions and Iran disclosure obligations remain in place.

First, the U.S. trade embargo against Iran remains in effect.77 Under these primary sanctions, U.S. persons and firms remain prohibited from engaging in any transactions or dealings directly or indirectly with Iran, including the indirect export of goods or services to Iran, unless otherwise authorized. U.S. and non-U.S. persons are also prohibited from evading U.S. sanctions or “causing” a sanctions violation by a U.S. person, such as by stripping or omitting information from transaction documents involving Iran.

Second, both U.S. and non-U.S. persons continue to face restrictions on the export and re-export to Iran of U.S.-origin goods or technology controlled under the Export Administration Regulations or International Trafficking in Arms Regulations.

Third, the United States retains various authorities to impose sanctions in response to Iran’s support for terrorism (e.g., Iran will remain designated as a State Sponsor of Terrorism under various statutes), human rights abuses, proliferation of weapons of mass destruction and their means of delivery, and destabilizing regional policies, such as in Syria and Yemen. The Government of Iran, including the Iranian military and intelligence establishment, remains sanctioned under these authorities.
Fourth, non-U.S. persons remain exposed to U.S. secondary sanctions for engaging in or facilitating transactions with persons or entities that will remain on the SDN List, unless such entities are on the Executive Order 13,599 List by virtue of their affiliation with the Government of Iran. SDN designees include the Iranian Revolutionary Guard Corps, an arm of the Iranian military with a pervasive presence in the Iranian economy, and other persons and entities engaged in conventional weapons proliferation or support for terrorism.

Fifth, so-called Section 219 SEC disclosure requirements under the Iran Threat Reduction Act remain in place following sanctions relief. Section 219 does not prohibit any specific conduct, but instead requires that “issuers” under the Securities Exchange Act of 1934 disclose in reports filed with the SEC various types of transactions in Iran undertaken by the issuer or its “affiliates.”

D. Other Significant Developments

1. Sudan

On January 13, 2017, OFAC issued a general license that authorizes all transactions with Sudan that were previously prohibited by the Sudanese Sanctions Regulations (SSR), 31 C.F.R. Part 538, and Executive Orders 13,067 and 13,412. President Obama also issued an executive order that would revoke the sanctions provisions applicable to Sudan through the previous executive orders on July 12, 2017, provided that the government of Sudan continues its cooperation with the United States on specified U.S. foreign policy priorities, including the cessation of hostilities in Sudanese conflict areas, improved humanitarian access in Sudan, and cooperation with the United States on addressing regional conflicts and the threat of terrorism.

The measures effectively suspend Sudan’s current status as a sanctioned country. U.S. persons are now able to conduct transactions with Sudan, and the property of the government of Sudan is unblocked.

The unique framework of this action means that the Trump Administration, based on the assessment and recommendation of the secretary of state and other incoming administration officials, will determine whether Sudan sanctions should be lifted permanently in July 2017.

For additional details about the January 2017 action with respect to Sudan, refer to our Client Alert: Sudan Embargo Lifted.

2. North Korea

North Korea has been at the center of worldwide attention in 2016 and 2017 due to its continued bellicosity, including underground nuclear testing and its launch of a new type of intermediate-range ballistic missile that it claimed could be nuclear-equipped. In response, the Trump Administration has rejected direct talks with North Korea and appears to be directing its attention at China. As discussed below, a key tool available to the United States are “secondary sanctions” that permit U.S. action against foreign companies that deal with North Korea. China, including its financial sector, would be a likely target of any U.S. secondary sanctions.

On February 18, 2016, President Obama signed the North Korea Sanctions and Policy Enhancement Act of 2016. Its scope was substantially broader than a contemporaneous United Nations Security Council Resolution in targeting North Korea’s weapons development and illicit financing activities and in creating new secondary sanctions against third-country firms that engage in certain types of activities in or with North Korea.

The Act provided for both mandatory and discretionary sanctions based on several criteria. The President must designate any firm or person determined to have knowingly assisted with the development of North Korea’s weapons of mass destruction program, delivery systems, or other military programs; exported luxury goods to North Korea; engaged in money laundering, counterfeiting or narcotics trafficking on
behalf of North Korea; engaged in cyberattacks on behalf of North Korea; or dealt in precious metals, minerals, or software related to weapons development. If a firm is designated under this “mandatory” subsection of the Act, then the President must impose asset-blocking requirements and prohibit all transactions in the property and property interests of that firm.

The President also has “discretionary” authority to select from a menu of sanctions to impose against persons who provide any material assistance to persons designated under UNSC resolutions, engage in bribery in North Korea, assist in the misappropriation of North Korean funds or financially support any of these activities. Those designated under this discretionary provision of the Act may be subject to one or more of the sanctions described in the Act, including the application of special measures for U.S. financial institutions to address money laundering; prohibitions on foreign exchange; prohibitions on transfers of credit or payments in or through the U.S. financial system; and certain other measures related to procurement, travel, and shipping.

In March 2016, President Obama signed Executive Order 13,722, blocking property of the Government of North Korea and the Workers’ Party of Korea, and authorizing additional designations of individuals or entities found to engage in certain activities including operation in the North Korean economy and certain dealings with the Government of North Korea or the Workers’ Party of Korea.82 OFAC simultaneously issued several general licenses authorizing certain activities (including noncommercial, personal remittances, support of nongovernment organizations, certain telecommunications activities, and the provision of certain legal services), but proceeded to make several rounds of new designations under this and other authorities targeting North Korea through 2016 and into 2017.83 For additional details, refer to our Client Alert: Enhanced North Korea Sanctions Adopted.

3. Burma

In October 2016, following a series of measures in 2016 and earlier easing of sanctions against Burma, President Obama signed an executive order terminating the national emergency with respect to Burma, revoking the Burma sanctions executive orders, and waiving other statutory blocking and financial sanctions on Burma.84 These measures ended the economic and financial sanctions against Burma.

4. Côte d’Ivoire

In September 2016, President Obama signed an executive order terminating the national emergency with respect to Côte d’Ivoire.85 President Bush, in 2006, had imposed sanctions against Côte d’Ivoire to address human rights abuses, political violence and unrest, and attacks against international peacekeeping forces. Such sanctions had been limited to the blocking of property and property interests of specifically identified individuals and entities. President Obama’s termination of the national emergency with respect to Côte d’Ivoire ended the sanctions program.

5. Kingpin Act

Since 2000, OFAC has designated almost 1,900 individuals and entities under the Kingpin Act for engaging in international narcotics trafficking.86 Throughout 2016 and early 2017, OFAC granted several general licenses, permitting transactions and activities that would otherwise be prohibited under the Kingpin Act. OFAC also designated certain individuals and institutions that have been connected to the Panama Papers leak.

In May 2016, OFAC published five general licenses, allowing transactions and activities related to the hotel at Millennium Plaza, Panama; two newspapers, La Estrella and El Siglo; the Soho Mall Panama; and transferring funds from Balboa Bank & Trust, seized by Panama. The general licenses for the Soho Mall Panama and Millennium Plaza are intended to assist with winding down transactions for a limited time; both entities are associated with the Waked Money Laundering Organization, which OFAC designated in May 2016. In June 2016, OFAC added general licenses for Importadora Maduro, S.A.;
6. Libya

In April 2016, President Obama signed Executive Order 13,726, blocking property and suspending entry into the United States of those contributing to the ongoing violence in Libya (including attacks by armed groups, human rights abuses, violations of the United Nations arms embargo, and misappropriation of Libyan resources). The order expanded the previous executive order from February 25, 2011, and blocks the property of any person responsible for or complicit in, directly or indirectly, any actions or policies that threaten peace, political transition, misappropriation of state assets, and other threats.
IV. Enforcement

Intense scrutiny in the areas of BSA/AML and sanctions enforcement is likely to continue based on current trends. At the start of 2017, all of the five largest U.S. banks by asset size had been subject to public regulatory actions relating to BSA/AML or sanctions deficiencies at some point within the past five years. Public disclosures also reflect that regulators and law enforcement remain active in these areas of enforcement, with a number of the largest financial institutions disclosing ongoing inquiries at the end of 2016.

In 2015, we saw record-setting penalties imposed, at times approaching and exceeding the billion-dollar mark. While the frequency and size of enforcement actions in 2016 was lighter than in the prior year, 2017 has seen a noticeable uptick in the announcement of significant enforcement actions that suggests the larger trend remains in the direction of very high regulatory expectations and continued enforcement.

A. AML Enforcement

Federal and state regulators announced two significant AML enforcement actions at the start of this year. Each action imposed large fines, eclipsing the biggest penalties assessed in 2016 and, consistent with current trends, required admissions of wrongdoing, significant AML program enhancements, and the appointment of independent monitors.

In January 2017, The Western Union Company forfeited $586 million to federal regulators and admitted to criminal violations that included willfully failing to maintain an effective AML program and allowing the processing of transactions in connection with an international consumer fraud scheme. Less than two weeks later, NYDFS announced a $425 million fine against Deutsche Bank AG and its New York branch in connection with violations of AML laws involving a long-running Russian “mirror trading” scheme. The actions reflect the culmination of a sustained focus by federal and state regulators on AML enforcement in recent years—a trend we expect to continue.

By contrast, 2016 was a year that saw a decrease in the number and size of public AML-related settlements compared to previous years, which had seen a number of record-setting fines and criminal prosecutions for violations of BSA/AML and sanctions laws. All indications continue to be that regulators and law enforcement agencies’ expectations remain high when it comes to the role that financial institutions are required to play in establishing effective compliance programs to monitor, detect, and report suspicious activity.

A variety of regulators, including DOJ, FinCEN, FINRA, the Office of the Comptroller of the Currency (OCC), SEC and NYDFS, brought notable enforcement actions over the past year. In connection with these actions, we observed several trends:

- Banking regulators continued to focus on the failure to file timely and effective SARs and, at the state level, NYDFS imposed the largest AML penalties of 2016, with a particular focus on foreign banks.

- AML actions involving broker-dealers were on the rise in 2016. These actions included the SEC’s first enforcement action premised entirely on a broker-dealer’s failure to file SARs, as well as FINRA’s imposition of the largest-ever AML fine. Microcap trading also continues to give rise to significant penalties.

- The imposition of independent monitors and consultants also continues to be a very common requirement in BSA/AML-related settlements. Last year alone, the oversight of at least a half-dozen independent parties was imposed as a condition of settlement by FinCEN, the SEC and NYDFS, including in actions involving banks, broker-dealers and gaming companies.
NYDFS imposed the most monitors last year. Acting under the newly appointed leadership of Superintendent Vullo, NYDFS required oversight by independent monitors, consultants or other third parties in all but one of the five BSA/AML enforcement actions NYDFS announced last year. The trend continued into 2017, with NYDFS requiring Deutsche Bank to engage an independent monitor for a two-year term. The common practice includes reserving the option to extend the monitorship at the end of the term, depending on the outcome of the institution’s remedial efforts.

FinCEN increased scrutiny of gaming companies, announcing three significant actions in 2016.

Compliance professionals continue to face independent liability for policing internal conduct as regulators advance efforts to hold individuals liable for compliance errors at their companies. FinCEN secured a significant legal victory at the start of 2016 when a federal district court in Minnesota held that the BSA permits it to bring suit against individuals for willfully violating the BSA’s AML program requirement.89

In 2016, the OCC, the SEC and FINRA each imposed AML-related fines against senior compliance officers and a CEO, ranging from $2,500 to $50,000. These fines were imposed on individuals in parallel with related AML settlements with those individuals’ financial institutions involving substantially larger penalties.

A brief summary of notable AML enforcement trends and the related actions that underlie them follows.

1. **Money Services Businesses**

Recently, federal regulators have focused their attention on AML compliance in the money services business (MSB) sector. In January 2017, Western Union agreed to forfeit $586 million and enter into agreements with DOJ, the Federal Trade Commission (FTC), several U.S. Attorneys’ Offices, and FinCEN.90 Western Union entered into a Deferred Prosecution Agreement (DPA), admitting to criminal violations including willfully failing to maintain an effective AML program and aiding and abetting wire fraud. According to the DPA, Western Union failed to file SARs identifying its own agents as suspicious actors who repeatedly facilitated consumer fraud-related transactions.91 Western Union also acquired a significant agent, FEXCO, that it knew, prior to the acquisition, had an ineffective AML program and had contracted with other agents who were facilitating significant levels of consumer fraud. Despite this knowledge, Western Union moved forward with the acquisition and did not remedy the AML failures or terminate the high-fraud agents.92 As part of the settlement with FinCEN, Western Union agreed to implement stricter AML/fraud policies. The FTC’s order required the appointment of an independent compliance auditor to assess compliance and issue periodic reports for a term of three years. The $586 million penalty will be used to reimburse consumers who were victims of the fraud.

2. **Banking Cases**

a) **NYDFS Continues to Impose Substantial AML Penalties**

At the state level, New York continues to impose substantial penalties for AML violations. In January 2017, NYDFS entered a consent order with Deutsche Bank in connection with a Russian mirror-trading scheme that went undetected among the bank’s Moscow, London and New York offices and allowed an estimated $10 billion to be moved out of Russia.93 The scheme involved closely related parties making a series of offsetting stock trades that, according to NYDFS, “lacked economic purpose and could have been used to facilitate money laundering or enable other illicit conduct.”94 Deutsche Bank cooperated with the investigation and agreed to pay a $425 million fine and hire an independent monitor for a two-year term in connection with violations of New York anti-money laundering law. NYDFS coordinated its investigation with the UK Financial Conduct Authority, which in parallel assessed a penalty of £168 million (approximately $210 million). NYDFS also used its announcement of the settlement to highlight the importance of its new risk-based anti-terrorism and anti-money laundering regulation, which became effective on January 1, 2017, and is described in more detail above under Regulatory Trends and Developments.
All of NYDFS’s public AML enforcement actions in 2016 focused on international banks. In December 2016, NYDFS fined Intesa Sanpaolo SpA and its New York branch $235 million and extended the term of engagement with its current independent consultant by up to two years for violations of New York anti-money laundering and Bank Secrecy Act laws. Intesa’s violations included a failure to follow written procedures to clear transactions and improperly closing alerts as “false positives” without investigation using ad hoc procedures when 41 percent of the alerts closed should have merited investigation. NYDFS also found that Intesa deliberately concealed information from bank examiners, and that from approximately 2002 to 2006 it used opaque methods and practices to conduct more than 2,700 U.S. dollar-clearing transactions, amounting to more than $11 billion, on behalf of Iranian clients and other entities possibly subject to U.S. economic sanctions. The consent order requires Intesa to strengthen compliance procedures, submit a revised compliance plan, and extend the engagement of its independent consultant for up to two years to test the results of remediation efforts.

In November 2016, NYDFS fined the Agricultural Bank of China $215 million for engaging in “intentional wrongdoing” resulting in AML violations. The NYDFS investigation discovered that the Bank ramped up its dollar-clearing activities through foreign correspondent accounts starting in 2013, even though NYDFS warned it not to increase those transactions until it significantly upgraded its internal compliance program. The Bank also severely curtailed the independence of the chief compliance officer at the New York branch, who tried to raise serious concerns to branch management and conduct internal investigations regarding suspicious activity. In addition to paying the penalty, the Bank agreed to take immediate steps to improve its legal compliance, including hiring an outside monitor.

In August 2016, NYDFS fined Mega International Bank of Taiwan (Mega Bank) $180 million, imposed a two-year independent monitor, and mandated reforms for AML violations. NYDFS found that the New York branch’s BSA/AML officer and chief compliance officer both lacked familiarity with U.S. regulatory requirements. In addition, the chief compliance officer had conflicted interests because she had key business and operational responsibilities along with her compliance role. Compliance staff at both the head office and branch failed to periodically review surveillance filter criteria designed to detect suspicious transactions. Also, numerous documents relied upon in transaction monitoring were not translated to English from Chinese, precluding effective examination by regulators. Finally, the New York branch procedures provided virtually no guidance concerning the reporting of continuing suspicious activities, had inconsistent compliance policies, and failed to determine whether foreign affiliates had adequate AML controls in place.

In the first quarter of 2016, NYDFS also announced entering into written agreements without penalties with the Industrial Bank of Korea and the National Bank of Pakistan, under which the banks agreed to enhance AML compliance and controls after examinations revealed significant deficiencies relating to risk management and compliance with relevant AML and sanctions laws. In the case of National Bank of Pakistan, the engagement of an independent third party was required for the purpose of conducting a transactional review. The Federal Reserve was also a party to both written agreements.

b) Continued Focus on Failures to File Timely and Effective SARs

2016 also saw a continued enforcement focus on the failure to maintain adequate anti-money laundering programs that result in the filing of timely and effective SARs. Among other examples, actions against Gibraltar Private Bank & Trust and Stearns Bank are illustrative of the types of conduct that have given rise to recent enforcement actions by FinCEN and the OCC.

In February 2016, FinCEN fined Gibraltar Private Bank & Trust $4 million for admitted failures in monitoring and detecting suspicious activity despite red flags. Gibraltar’s transaction monitoring system contained incomplete and inaccurate account opening information and customer risk profiles, and its automated monitoring system generated an unmanageable number of alerts, including large numbers of false positives. Gibraltar also failed to properly train its compliance staff, and failed to develop and implement an adequate Customer Identification Program. As a result, Gibraltar failed to timely file at least 120 SARs involving nearly $558 million in transactions occurring from 2009 to 2013. Gibraltar also unreasonably delayed SAR reporting regarding accounts related to a $1.2 billion Ponzi scheme led by
Florida attorney Scott Rothstein. In parallel, OCC fined Gibraltar $2.5 million, payment of which offset FinCEN’s $4 million fine.  

In April 2016, the OCC settled its own SAR timeliness enforcement with Stearns Bank. Beginning in March 2010, Stearns Bank became aware of suspicious transactions associated with the manipulation and fabrication of accounts receivables and factoring invoices, but failed to follow its internal policies and procedures governing the monitoring and reporting of suspicious activity, including the maintenance of appropriate documentation to support its SAR determinations. As a result, Stearns Bank failed to file timely SARs and was required to pay a $1 million penalty.  

On February 27, 2017, FinCEN announced a $7 million civil penalty assessment against Merchants Bank of California for failing to establish an adequate AML compliance program, conduct due diligence on foreign correspondent accounts, and detect and report suspicious activity. The California-based community bank provided banking services to MSBs such as check-cashers and money transmitters, but did not adequately assess the money laundering risks presented by servicing the MSBs. In a press release announcing the settlement, FinCEN Acting Director Jamal El-Hindi described the bank as “an institution run by insiders essentially to provide banking services to MSBs that the insiders owned, combined with directions from Bank leadership to staff to ignore BSA requirements with respect to those MSB customers and others.” The bank’s leadership, according to FinCEN’s assessment, impeded the investigation and reporting of suspicious activity by threatening employees who attempted to report suspicious transactions in accounts affiliated with bank executives. Among other violations, the bank also failed to conduct independent testing commensurate with its customer complexity and risk profile and failed to designate a BSA officer. Concurrent with the FinCEN order, the OCC imposed a $1 million penalty (which FinCEN credited toward satisfying its own $7 million penalty) for AML compliance failures that led to violations of previous OCC consent orders.  

3. Securities Cases
   a) Microcap Trading Continues to Give Rise to AML Actions

FINRA continues to focus on AML violations relating to microcap trading. In December 2016, FINRA fined Credit Suisse Securities (USA) LLC $16.5 million for AML compliance program failures, alleging that Credit Suisse failed to effectively review microcap trading for AML purposes from January 2011 to September 2013 and to review potentially suspicious money transfers from 2011 to 2015. FINRA concluded that Credit Suisse’s systems and procedures were not adequately designed to detect potentially suspicious transactions in order to cause the filing of SARs. Although the firm used an automated surveillance system to identify red flags, it failed to properly implement the system, including by failing to input adequate data and to use applicable risk scenarios in its assessments. FINRA also alleged that Credit Suisse did not have adequate staffing to review the tens of thousands of alerts its automated system generated in any given year.

Another major microcap-related enforcement action was resolved in 2015 against Oppenheimer & Co. There, the SEC and FinCEN fined the broker-dealer a total of $20 million for failing to establish and implement an adequate anti-money laundering program, failing to conduct adequate due diligence on a foreign correspondent account, and failing to comply with requirements under Section 311 of the USA PATRIOT Act. Sixteen customers engaged in patterns of suspicious microcap trading through Oppenheimer’s branch offices. Oppenheimer failed to report patterns of activity in which customers deposited large blocks of unregistered or illiquid microcaps, moved large volumes of microcaps among accounts with no apparent purpose, or immediately liquidated those securities and wired the proceeds out of the account. In addition, Oppenheimer designated a customer foreign financial institution as “high risk” but failed to assess the specific risks as a foreign financial institution or conduct adequate due diligence. As a result, Oppenheimer did not detect or investigate numerous suspicious transactions conducted through the account, including prohibited third-party activity and illegal penny stock trading.
b) **FINRA Imposes Its Largest-Ever AML Fine**

In May 2016, FINRA issued its largest-ever AML fine totaling $17 million against two Raymond James entities – Raymond James & Associates, Inc., and Raymond James Financial Services, Inc. (collectively, Raymond James).\(^{108}\) FINRA alleged that Raymond James’ significant growth between 2006 and 2014 was not matched by commensurate growth in their AML compliance systems and processes. This left the firms unable to establish AML programs tailored to their businesses, and forced them instead to rely on a patchwork of written procedures and systems across different departments to detect suspicious activity. The end result was that red flags of potentially suspicious activity went undetected or inadequately investigated. FINRA stated that these alleged failures were “particularly concerning” because Raymond James was previously sanctioned in 2012 for inadequate AML procedures and, as part of that settlement, had agreed to review its program and procedures and certify that they were reasonably designed to achieve compliance.

c) **SEC’s First Stand-Alone Action Premised on Failure to File SARs**

The SEC’s June 2016 settlement with Albert Fried & Company (AF & Co.) is significant in that it represents the first SEC enforcement action premised solely on a broker-dealer’s failure to file SARs and not any other securities law violations.\(^{109}\) AF & Co., a Wall Street-based brokerage firm, agreed to pay a $300,000 penalty to settle charges that it failed to file SARs for more than five years despite red flags tied to its customers’ high-volume liquidations of low-priced securities. On more than one occasion, an AF & Co. customer’s trading in a security on a given day exceeded 80 percent of the overall market volume. In other instances, customers were trading in stocks issued by companies that were delinquent in their regulatory filings or involved in questionable penny stock promotional campaigns. The SEC found that AF & Co. violated Section 17(a) of the Securities Exchange Act of 1934 and Rule 17a-8 thereunder.

4. **Gaming Companies**

Over the past year, FinCEN increased its scrutiny of gaming companies, giving rise to three actions, summarized below.

In October 2016, FinCEN assessed a $12 million civil money penalty against Cantor Gaming.\(^{110}\) Concurrently, Cantor Gaming reached a non-prosecution agreement with the U.S. Attorney’s Office for the Eastern District of New York, including a fine of $10.5 million and forfeiture of $6 million.\(^{111}\) FinCEN determined that Cantor Gaming failed to implement and maintain an effective AML program by failing to have sufficient internal controls and mandatory independent audits and failing to sufficiently train its officers and employees. FinCEN determined that Cantor Gaming also failed to properly and timely file Currency Transaction Reports and SARs and committed thousands of recordkeeping violations, including by failing to keep required records on its highest-volume patron, who placed more than $300 million in wagers between 2010 and 2013.

In July 2016, FinCEN assessed a $2.8 million civil money penalty against Hawaiian Gardens Casino for violations of its BSA/AML program and reporting obligations.\(^{112}\) IRS investigations in 2011 and 2014 revealed multiple BSA violations, and many of the violations uncovered in 2011 remained unaddressed in 2014, despite identification by the casino’s independent consultant in 2013. Hawaiian Gardens failed to implement and maintain an effective AML program, failed to report large cash transactions, failed to file SARs and failed to keep certain required records. Although the casino had a variety of tools including casino surveillance and open source information at its disposal, it failed to use those tools to gather customer information and, as a result, 80 percent of the SARs it filed between January 2013 and September 2014 included at least one unknown subject. Certain employees had also been identified multiple times by the IRS for assisting customers with structuring. Similarly, Hawaiian Gardens continued to conduct business with patrons that they had identified as suspicious, even after they had repeatedly refused to provide identification information.

In April 2016, FinCEN fined Sparks Nugget $1 million for violating AML program requirements, reporting obligations, and recordkeeping requirements.\(^{113}\) Sparks Nugget allegedly lacked a “culture of
compliance”: the casino routinely disregarded its BSA compliance manager, chose not to file SARs that the compliance manager believed should have been filed, instructed her to not interact with the IRS’s BSA auditors, and prevented her from reviewing a copy of the IRS’s completed exam report. The management committee that Sparks Nugget established to determine whether to file SARs never met, and some committee members were unaware that they were even on the committee. Further, Sparks Nugget lacked any mechanism to document decisions not to file SARs, and its day-to-day managers maintained that no suspicious activity ever transpired at the casino.

5. High Risk of Personal Liability for Compliance Officers

The start of 2016 saw a significant legal development in FinCEN’s landmark Haider case: A federal district court held that the Bank Secrecy Act permits FinCEN to bring suit against individuals for willfully violating the BSA’s AML program requirement. The decision was significant not only because it appears to be the first court decision in a FinCEN civil action but also because the opinion addressed important aspects of FinCEN’s enforcement authority. The case arose after MoneyGram settled with DOJ for admitted AML program failures, and Thomas Haider, former chief compliance officer of MoneyGram, was fined $1 million by FinCEN for failing to ensure an effective AML program and failing to ensure timely filing of SARs. In addition to the $1 million penalty, FinCEN also sought to bar Mr. Haider from the financial industry. Haider, however, sought to dismiss the case, arguing that he could not be held liable under 31 U.S.C § 5318(h) because that section of the BSA—which requires each regulated “financial institution” to establish an AML program—applies only to financial institutions, rather than individuals. The court rejected Haider’s argument, concluding that his focus on the text of § 5318(h) was misplaced. Instead, the court looked to the BSA’s general civil penalty provision in § 5321(a), which permits FinCEN to assess civil penalties against individuals for willful violations of the BSA, with the exception of two BSA sections. The court held that because § 5318(h) is not one of those two exceptions, FinCEN can assess civil money penalties against financial institution officers or employees who violate § 5318(h). The court’s analysis of this issue is not expansive, and its reasoning—which arguably expands the substantive requirements of the BSA based on the civil penalty provision—is not free from doubt. The Haider opinion is unlikely to be the last word on the issue of whether individuals may be held liable under the BSA for what appear to have been, at bottom, institutional failures.

Regulators are showing a sustained interest in holding compliance officers personally liable for compliance deficiencies. Last year, Raymond James’ former AML compliance officer Linda Busby was fined $25,000 and suspended for three months for alleged failures to detect or investigate red flags. In March 2016, the OCC fined Charles Sanders, the former chief compliance officer and chief risk officer of Gibraltar Private Bank & Trust, $2,500 for failing to “file suspicious activity reports on a set of accounts for a customer that was later convicted of crimes related to an illegal Ponzi scheme.” Sanders was also ordered to disclose the settlement to any future employers that fall under the definition of a “depository institution.” Finally, in mid-October 2016, the SEC settled an action against Lia Yaffar-Pena, the former president and CEO of Miami-based brokerage firm E.S. Financial, for aiding and abetting and causing violations of AML rules by allowing foreign entities to buy and sell securities without verifying the identities of the non-U.S. citizens who beneficially owned them. Yaffar-Pena agreed to a one-year supervisory suspension and payment of a $50,000 penalty. The SEC had previously settled an enforcement action against Yaffar-Pena’s brokerage firm, E.S. Financial, for $1 million for the same alleged violations.

While the unmistakable trend in recent years has been an increasing risk of individual liability for compliance professionals, it remains to be seen how the new Administration’s policymakers will balance the desire for individual accountability against the potential negative effects that such actions may have on the ability of the industry to hire and retain qualified professionals willing to take on the job. Compliance professionals, who are an important backstop and internal check on the business, are now facing the specter of being held individually responsible after the fact for the advice they provide on where lines can be drawn in today’s complex regulatory environment.
B. Sanctions Enforcement

Last year was a relatively light year for OFAC enforcement, both in terms of the number of cases finalized and the total penalty amounts paid in those cases. In 2016, OFAC announced penalties or settlements in just nine cases, with a total penalty/settlement amount of just over $21 million. By contrast, 2015 saw 15 cases totaling nearly $600 million, and 2014 saw 23 cases totaling $1.2 billion. Much of the penalty and settlement total amounts in the preceding year was the result of major enforcement actions against global financial institutions, such as the $963 million settlement with BNP Paribas SA in 2014, the $329 million settlement with Crédit Agricole Corporate and Investment Bank in 2015, and the $258 million settlement with Commerzbank AG in 2015.

The largest sanctions settlement in 2016 was in July for just $7.6 million and with a non-financial institution. That case involved a Texas company, Alcon Laboratories, Inc., and its Swiss affiliate, who agreed to settle with OFAC and the Commerce Department’s Bureau of Industry and Security (BIS) for potential civil liability arising from sales and exports of medical end-use surgical and pharmaceutical products to distributors in Iran and Sudan without OFAC authorization. Among the aggravating factors under OFAC’s Enforcement Guidelines was the fact that the company’s senior management knew of the conduct giving rise to the apparent violations and that the company “is a sophisticated multinational corporation with extensive experience in international trade.”

The next-largest settlement in 2016 was in November, when National Oilwell Varco, Inc., and its subsidiaries agreed to set their potential civil liability for apparent violations of U.S. sanctions against Cuba, Iran, and Sudan. OFAC alleged that the company had approved commission payments by its subsidiary to a UK-based entity related to the sale and export of goods to Iran and had engaged in other transactions involving the sale and export of goods to Iran, Cuba, and Sudan. The case was an egregious one under OFAC’s Enforcement Guidelines because, among other reasons, the company’s senior-level finance executives approved the commission payments and the company appeared to have willfully blinded itself by “acquiescing” to its subsidiary’s “deliberate non-identification of Iran in its communications.”

The largest OFAC settlement with a financial institution in 2016 was in February, when Barclays Bank Plc agreed to remit $2.48 million to settle its potential civil liability for 159 alleged violations of U.S. sanctions against Zimbabwe. OFAC alleged that the bank processed over $3 million in transactions to or through financial institutions in the United States, including its New York branch, for or on behalf of corporate customers of Barclays Bank of Zimbabwe Limited that were owned 50 percent or more by a person on OFAC’s SDN List. OFAC found that the Zimbabwe bank’s KYC procedures “were ambiguous and difficult to follow with respect to the requirement to identify related parties and/or beneficial owners of corporate customers.” OFAC determined that the alleged violations constituted a “non-egregious case,” and recognized as mitigating factors Barclays’ significant remedial actions and substantial cooperation.

No 2016 action met the level in 2015, where we saw two major sanctions-related enforcements. First, in March 2015, Commerzbank AG settled with DOJ, OFAC, the Federal Reserve Board of Governors (FRB), and NYDFS for $1.4 billion for violations of the International Emergency Economic Powers Act (IEEPA) and for failing to have an effective AML program, failing to conduct due diligence on foreign correspondent accounts, and failing to file SARs. Regarding sanctions, Commerzbank concealed hundreds of millions of dollars in transactions on behalf of sanctioned Iranian and Sudanese businesses, even though managers inside the bank raised red flags about its sanctions-violating practices. Commerzbank also admitted to AML deficiencies that made it a conduit for over a billion dollars of the Olympus securities fraud. In addition to the fine, Commerzbank also agreed to implement rigorous internal controls. Second, in October 2015, Crédit Agricole paid $787.3 million to settle with DOJ, OFAC, FRB, and NYDFS over violations of the IEEPA and the Trading with the Enemy Act. Between August 2003 and September 2008, subsidiaries of the Bank in Geneva knowingly and willfully moved approximately $312 million through the U.S. financial system on behalf of sanctioned entities located in Sudan, Burma, Iran and Cuba.
This decrease in OFAC penalty cases during 2016 is unlikely to represent a change or a significant shift in enforcement policy. To the contrary, the Trump Administration is expected to maintain or even intensify sanctions enforcement, especially with respect to U.S. sanctions against Iran. And civil penalties, when they are imposed, will now be greater than before: On February 9, 2017, OFAC issued regulations to implement the Federal Civil Penalties Adjustment Act of 1990, adjusting for inflation the maximum amount of civil monetary penalties that may be assessed.126 Violations of the IEEPA, the statute under which most OFAC sanctions regulations have been promulgated, may now be subject to a penalty of the greater of $289,238 (per violation) or twice the amount of the underlying transaction.

In addition to three penalties or settlement agreements in January 2017 alone, the new year has also delivered the largest OFAC penalty ever against a non-financial institution. On March 7, 2017, OFAC, BIS, and DOJ announced a $1.2 billion settlement agreement with ZTE Corp. for civil and criminal violations of the export control and sanctions laws, which included a guilty plea for conspiring to violate the IEEPA. Of the settlement amount, nearly $101 million will go to OFAC for violations of U.S. sanctions against Iran. The settlement agreement followed an investigation into ZTE Corp.’s “multi-year and systematic practice of utilizing third-party companies to surreptitiously supply Iran with a substantial volume of U.S.-origin goods, including controlled goods appearing on the Commerce Control List (CCL).” ZTE’s misconduct also included serious charges of obstructing justice and taking affirmative steps to mislead the U.S. government, which likely influenced the record-setting penalty.127
WilmerHale’s interdisciplinary AML and Economic Sanctions Compliance and Enforcement Group brings together leading practitioners to focus on our clients’ most challenging AML- and economic sanctions–related regulatory, examination and enforcement issues. The team has a wealth of knowledge and government experience at the forefront of AML and sanctions policy and enforcement. Our lawyers have worked in the U.S. Department of Justice, U.S. Attorneys’ Offices, the U.S. Department of the Treasury, the U.S. Department of State, the Central Intelligence Agency and the National Security Agency, the Securities and Exchange Commission, the Office of the Comptroller of the Currency, the White House, and the United States Congress. This depth of experience enables us to assist clients in anticipating and understanding the government’s priorities, communicating with regulators and key stakeholders, and resolving their most challenging matters and law enforcement proceedings.

**Regulatory:** We advise financial institutions on a complex array of regulations issued by the Financial Crimes Enforcement Network, the Office of Foreign Assets Control, and state and federal banking and securities supervisors. We assist clients in preparing for and responding to regulatory examinations conducted by banking and securities regulators. Our attorneys draft regulatory comment letters and advise financial institutions and trade associations on the implications of forthcoming rule-makings. We also advocate for our clients regarding regulatory and statutory issues in Congress with key oversight and policymaking committees.

**Compliance:** We provide compliance training, advise on strategic and tactical compliance matters, and assist our clients in drafting policies and procedures to enhance their compliance programs. We help many U.S. and non-U.S. clients develop and implement internal policies and procedures to promote compliance with applicable AML and sanctions requirements, which often present complex challenges for financial institutions with global operations. Our advice includes corporate compliance programs, contractual assurances, technology control and vendor management plans, transaction and customer screening, and in-house training and compliance reviews.

**Enforcement:** We represent a diverse array of foreign and domestic financial institutions that have found themselves the targets of enforcement actions by federal and state regulators and of congressional inquiries. Our experience spans the lifecycle of enforcement, from responding to initial formal and informal requests for information through negotiating consent orders and compliance with consent orders. We also represent financial institutions in federal and state criminal investigations and frequently advise clients on matters involving voluntary self-disclosures of sanctions violations. Our attorneys have assisted financial and other institutions with their responses to nearly all of the major congressional inquiries regarding AML issues over the past two decades.

**Transactional Counseling:** AML and sanctions compliance issues arise in a variety of business transactions, including mergers and acquisitions, joint ventures, trade financing, and other specialized transactions. WilmerHale has extensive experience counseling financial firms on AML- and OFAC-related transactional issues. We work with colleagues in our Corporate Practice and Transactional Department to review and assess the risks associated with potential transactions, and advise on the allocation of risks and liabilities between the parties. Where appropriate, we design potential remediation.
FOR MORE INFORMATION ON AML AND SANCTIONS MATTERS, PLEASE CONTACT:

New York

Sharon Cohen Levin +1 212 230 8804 sharon.levin@wilmerhale.com
Boyd M. Johnson III +1 212 230 8862 boyd.johnson@wilmerhale.com

Washington DC

Franca Harris Gutierrez +1 202 663 6557 franca.gutierrez@wilmerhale.com
Reginald J. Brown +1 202 663 6430 reginald.brown@wilmerhale.com
Ronald I. Meltzer +1 202 663 6389 ronald.meltzer@wilmerhale.com
Michael J. Leotta +1 202 663 6526 michael.leotta@wilmerhale.com
Jeremy Dresner +1 202 663 6176 jeremy.dresner@wilmerhale.com
David M. Horn +1 202 663 6749 david.horn@wilmerhale.com


9 See, e.g., In re Brown Brothers Harriman & Co., FINRA Letter of Acceptance, Waiver, and Consent, No. 2013035821401, at 2 (Feb. 4, 2014) (“BBH was obligated under federal law to investigate customer activity on a risk basis; omnibus accounts transacting in higher-risk activity, such as suspicious penny stock transactions, merited additional scrutiny.”), http://disciplinaryactions.finra.org/Search/ViewDocument/35225.


14 Id. at 2.

15 See id. at 4-5 for a full list of red flags.


17 The Federal Banking Agencies are the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation, the National Credit Union Administration, and the Office of the Comptroller of the Currency.


21 Hearing on the Nomination of Steven Mnuchin to be Secretary of the Treasury, Senate Finance Committee, Written Responses to Questions for the Record, at 29, 93 (Jan. 2017), https://dbjbjzgnk95t.cloudfront.net/0884000/884398/mnuchinresponses.pdf.


23 Maria Vullo, NYDFS Superintendent, Remarks Delivered at event hosted by the Exchequer Club (Jan. 18, 2017).


26 NYDFS Superintendent’s Regulations § 504.1.

27 See N.Y. Fin. Serv. Law § 301(c)(4) (giving NYDFS the authority to refer matters to the New York Attorney General); N.Y. Banking Law § 672 (prohibiting the falsification of books, reports, or statements of banks).


29 Covered entities have until March 1, 2018, to comply with the following provisions: 500.04(b) (annual CISO reports to the board of directors), 500.05 (penetration and vulnerability assessments), 500.09 (risk assessments), 500.12 (multi-factor authentication) and 500.14(b) (cybersecurity awareness training). Covered entities have until September 1, 2018, to comply with the following provisions: 500.06 (audit trails), 500.08 (application security), 500.13 (data retention and destruction), 500.14 (a) (monitoring of authorized users) and 500.15 (encryption). Covered entities have until March 1, 2019, to comply with the third-party service provider provision, section 500.11.

30 Board of Governors of the Federal Reserve System, Interagency Guidelines Establishing Information Security Standards, § II, https://www.federalreserve.gov/bankinforeg/interagencyguidelines.htm. See 23 NYCRR § 500.01(g) (defining “Nonpublic Information” as “all electronic information that is not Publicly Available Information and is: (1) Business related information of a Covered Entity the tampering with which, or unauthorized disclosure, access or use of which, would cause a material adverse impact to the business, operations or security of the Covered Entity; (2) Any information concerning an individual which because of name, number, personal mark, or other identifier can be used to identify such individual, in combination with any one or more of the following data elements: (i) Social Security number, (ii) driver’s license number or non-driver identification card number, (iii) account number, credit or debit card number, (iv) any security code, access code or password that would permit access to an individual’s financial account, or (v) biometric records; (3) Any information or data, except age or gender, in any form or medium created by or derived from a health care provider or an individual and that relates to (i) the past, present or future physical, mental or behavioral health or condition of any individual or a member of the individual’s family, (ii) the provision of health care to any individual, or (iii) payment for the provision of health care to any individual.”).


32 See SEC 2017 Exam Priorities at 4-5.

33 FINRA 2017 Exam Priorities, at 8.
The FATF conducted its on-site visit from January 18, 2016, to February 5, 2016.

FinCEN issued its final customer due diligence rule on May 11, 2016, after the close of the FATF on-site visit.


Id. at 3.


OFAC also designated two additional individuals—Evgeniy Mikhailovich Bogachev and Alexsey Alekseyevich Belan—for their cyber-enabled misappropriation of financial information and personal identifiers for private financial
gain. OFAC stated that Mr. Bogachev developed the Zeus malware, which is associated with the theft of financial information and other criminal activity. According to OFAC, Mr. Bogachev directly benefited from the use of the malware by other cybercriminals, and he also used a form of “ransomware” to hold at least 120,000 U.S. victims’ data hostage for financial gain in excess of $100 million. OFAC stated that Mr. Belan’s attacks on at least three U.S.-based e-commerce companies’ computer networks led to the theft of e-mail addresses, customer names and encrypted passwords, which he sold for private financial gain.


75 U.S. Dep’t of Treasury, Resource Center: List of Persons Identified as Blocked Solely Pursuant to Executive Order 13,599, https://www.treasury.gov/resource-center/sanctions/Programs/Pages/13599_list.aspx.


39


Liability for Apparent Violations of the Zimbabwe Sanctions Regulations


Id. at 1.


CFIUS Actions Highlight Focus on Critical Technologies and Infrastructure

June 3, 2016
By Ambassador Robert M. Kimmitt, Benjamin A. Powell, Stephen W. Preston, Jason C. Chipman

In its latest report to Congress, the Committee on Foreign Investment in the United States (CFIUS), the US government’s interagency body that vets foreign investment with national security implications, reported that it reviewed a record 147 notices in 2014, a 52% increase from 2013.¹ This data, as well as public information about more recent individual cases reviewed by CFIUS, underscores that CFIUS is taking a more expansive view of its jurisdiction and will scrutinize investments affecting US critical technologies and critical infrastructure. Parties considering transactions subject to CFIUS review should take account of this recent CFIUS precedent in conducting national security due diligence for their deals.

I. Report Highlights

The CFIUS Annual Report suggests several trends in its review of covered transactions:

- CFIUS reports a significant increase in the total number of cases filed in 2014 (147), compared with 2013 (97) and 2012 (114).
- The percentage of cases that CFIUS extended into an additional 45-day investigation period decreased, however, to 35% in 2014 from 49% in 2013. The high number of investigations in 2013 was due partly to the US government shutdown.
- Although no transaction was formally blocked in 2014, 12 notices were withdrawn during either the initial review or investigation stage. Some of these notices may have been withdrawn and not refiled as a result of CFIUS-related national security concerns, but others may have been withdrawn for unrelated commercial or regulatory reasons.
- CFIUS required mitigation in 9 cases in 2014, or 6% of all notices, a decrease from 11% of the CFIUS notices reviewed in 2013.

The report indicates a marked increase in transactions in the computer and electronic product sector (29 notices) in 2014, a 142% increase from 2013 (12 notices). As described further below, the largest number of 2014 notices occurred in the semiconductor and other electronic component manufacturing subsectors (12 notices). A significant share of notices also occurred in the professional, scientific, and technical services (14 notices) and utilities (13 notices) sectors.

Acquisitions by investors from Canada, China, Japan and the United Kingdom continued to account for the largest share of notices in 2014, with 15, 24, 10 and 21 notices, respectively. Over the 2012 to 2014 period, the top three home countries of foreign acquirers were China (68 notices), the United Kingdom (45 notices) and Canada (40 notices).
II. US Critical Technologies and Critical Infrastructure Concerns

The report and recent action by CFIUS—including publicized actions subsequent to the 2014 reporting period—suggest that CFIUS will closely scrutinize investments affecting US critical technologies and critical infrastructure and may interpret its jurisdiction expansively in cases of doubt.

As described below, for companies contemplating transactions that may be subject to CFIUS review, transaction diligence should be comprehensive and include: both direct US government contracts held by the US business and potential indirect or channel sales to government agencies via third-party distributors (to the extent ascertainable); possible past US government grants to foster the development of the subject technology; potential defense applications for any subject technology (even those not directly known by the parties); and proximity of US facilities to military installations or other sensitive infrastructure. Such diligence will assist parties in assessing the likelihood of CFIUS approval, addressing any national security issues proactively at an early stage of review, and developing appropriate risk-sharing provisions in contract documents.

A. CFIUS’s Growing Semiconductor and Defense Supply Chain Focus

CFIUS has reviewed a significant number of semiconductor deals in recent years, including 12 notices in 2014, 6 notices in 2013, and 12 notices in 2012. Over the past year, China-based investors have reportedly bid for at least 10 semiconductor businesses, most of which were in the United States. The US Department of Defense, one of the CFIUS agencies, has also created a special task force to examine semiconductor concerns for the defense supply chain.

This does not mean that CFIUS will not allow semiconductor deals to proceed. For example, in November 2015, NXP Semiconductors N.V. announced that it received clearance on the $1.8 billion sale of its RF Power business to Jianguang Asset Management Co. Ltd., a Chinese state-owned investment company. In November 2015, Integrated Silicon Solution Inc. received CFIUS approval for its $640 million deal with a China-backed investment group led by Hua Capital and Summitview Capital. In October 2015, OmniVision Technologies Inc., a camera sensor-maker, and a group including Chinese investors Hua Capital Management Co., Ltd., CITIC Capital Holdings Limited and GoldStone Investment Co., Ltd., received CFIUS approval for a $1.9 billion deal.

CFIUS concerns did apparently lead to the blocking in January 2016 of the $2.9 billion acquisition by a Chinese consortium led by GO Scale Capital of an 80% stake in Koninklijke Philips N.V.’s California-based Lumileds business, a supplier of semiconductor light-emitting diode (LED) lighting components for general illumination, consumer electronics and automotive applications. The transaction did not appear on its face to give rise to major national security concerns in light of Lumileds’ known customer base, and CFIUS did not describe the precise concerns underlying its opposition.

Reports indicate, however, that solid-state lighting, such as advanced LEDs and component materials, may have a variety of sophisticated defense applications. Given that CFIUS reportedly told the company...
that “it couldn't have been aware of the reason for its objections,” it is possible that Lumileds products' comprised part of the supply chain for sensitive military technology and/or benefitted directly or indirectly from US government development grants.4

The failure of the Lumileds deal—for which the parties had already committed an estimated €500 million in separation costs—underscores the need for comprehensive diligence for deals subject to CFIUS jurisdiction. Although companies may not always have insight into government end users for third-party sales, understanding the sensitivity and potential military applications of products may assist parties in developing informed CFIUS assessments and appropriate risk-sharing provisions in contract documents. Also, members of Congress have raised concerns with semiconductor and other related transactions, requiring companies to be proactive and potentially engage publicly to address national security concerns.5

B. CFIUS Interprets Its Jurisdiction Expansively

Particularly in sectors of increased CFIUS risk, some parties have attempted to structure deals to avoid CFIUS scrutiny, but CFIUS will interpret its jurisdiction expansively in cases of doubt. For example, Western Digital Corp., a US hard drive manufacturer, and Tsinghua Unisplendour Corp., a Chinese business, recently abandoned a $3.8 billion deal for Tsinghua's acquisition of a 15% stake in Western Digital after CFIUS indicated that it would review the transaction.6 The parties had been engaged in a months-long process with CFIUS to determine whether CFIUS even had jurisdiction to review the transaction.7 The parties reportedly structured the transaction to avoid implicating CFIUS jurisdiction, which requires acquisition by a foreign person of “control” over a “US business.”8 Although CFIUS has a safe harbor provision for purely passive investments,9 the deal would have reportedly made Tsinghua the largest Western Digital shareholder and granted Tsinghua a representative on the Western Digital board.10

C. CFIUS Scrutinizes Co-Location and Critical Infrastructure

Since CFIUS forced the divestment of an acquisition by Ralls Corporation, a US company owned by Chinese nationals, of a wind farm overlooking a US naval base in Oregon in 2012, co-location and critical infrastructure issues have become an increasingly important CFIUS review factor. In 2013, CFIUS reportedly scrutinized the acquisition of Smithfield Foods Inc., a major US pork producer, by a Chinese investor, Shuanghui International Holdings Ltd., in part over Smithfield's proximity to a variety of US military installations in Virginia.11 That year, CFIUS also forced Chinese energy firm Cnooc Ltd. to divest various oil platforms and approximately 200 leases in the Gulf of Mexico due to proximity to a US naval base in Louisiana as a condition to approving its acquisition of Nexen Inc., a Canadian energy firm.12 This year, negotiations between Terex Corp., a US construction and crane firm, and Zoomlion, a Chinese buyer, were reportedly complicated by Terex's US government contracts and operations at the Port of Long Beach, California, the second-largest US port, before Zoomlion backed out of negotiations in May.13
This recent CFIUS precedent suggests that parties should closely scrutinize the locations of all US offices and facilities associated with the US business of the target company. In the case of multiple US facilities, CFIUS may require security procedures, assurances or possible divestment of the assets in question.


8 See 31 C.F.R. §§ 800.207, 800.301.

9 31 C.F.R. § 800.302(b).

10 Similarly, in 2011, CFIUS forced Huawei to divest its purchase of US technology firm 3Leaf, even though it purchased certain aspects of 3Leaf's patents and servers out of bankruptcy. Given 3Leaf's financial condition, the parties apparently expected that the deal would not be reviewable by CFIUS. Shayndi Raice, "Huawei Set Back on Deal in U.S.,” The Wall Street Journal, Feb. 15, 2011.


4 | CFIUS ACTIONS HIGHLIGHT FOCUS ON CRITICAL TECHNOLOGIES AND INFRASTRUCTURE
President Trump Issues Cybersecurity Executive Order

May 12, 2017
By Benjamin A. Powell, Jonathan G. Cedarbaum, Leah Schloss

On May 11, President Trump signed his long-awaited Executive Order on “Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure.” Much of the Order mandates efforts to improve the government’s own information technology (IT) and cybersecurity practices, but several directives focus on the private sector: (i) a report to the President within 90 days on whether publicly-traded companies operating critical infrastructure should have to make fuller public disclosures concerning their cybersecurity practices; (ii) a report to the President within 90 days on the electric sector’s ability to respond to and mitigate an attack leading to a prolonged outage; (iii) a report to the President within 90 days on cyber threats confronting defense industrial base (DIB) companies and their supply chains; and (iv) a requirement that agencies conform their cybersecurity guidance documents to the National Institute of Standards and Technology’s (NIST) Cybersecurity Framework, which may lead to new requirements for government contractors. The Order also calls for reports on the government’s ability to assist and defend companies operating critical infrastructure systems at greatest risk, deterring cyber threats from abroad, building more effective defenses against botnets, international cybersecurity priorities, and cyber workforce development.

Cybersecurity of Critical Infrastructure

To support the risk management efforts of the owners and operators of critical infrastructure (CI), the Order requires various designated agencies to report to the President, within 90 days:

- Whether Federal policies and practices are sufficient “to promote appropriate market transparency of cybersecurity risk management practices” by CI entities, focusing on publicly-traded CI entities. This suggests that the Trump Administration may be exploring the possibility of creating new requirements for companies, particularly those that are publicly-traded, to report their cybersecurity risk management practices to investors. In the event that this assessment leads to more extensive requirements or more aggressive enforcement by the Securities and Exchange Commission, publicly-traded companies should consider reviewing their securities disclosures to confirm that cybersecurity risks and risk management practices are appropriately disclosed to investors.

- Regarding the potential scope and duration of a prolonged power outage associated with a significant cyber incident, the country’s readiness to manage the consequences of such an incident, and any gaps or shortcomings in assets or capabilities. The specific focus on the electric sector may signal cybersecurity priorities for the Trump Administration going forward, and companies in this sector in particular should seek opportunities to engage with Administration officials as they conduct the assessments and reporting required by the Order.

Attorney Advertising

WilmerHale is a limited liability partnership. WilmerHale's principal law offices: 600 Massachusetts Avenue, N.W., Washington, D.C. 20001; +1 202 659 6000. Our London office is operated under a separate limited liability partnership with registered foreign lawyer authorized and regulated by the Solicitors Regulation Authority (SRA No. 207648). Our professional rules can be found at www.wilmerhale.com/code-of-conduct.phtml. A list of partners and their professional qualifications is available for inspection at our UK office. In Brussels, we are represented by a foreign law firm representative office. This material is for general informational purposes only and does not represent or convey any advice or its any particular set of facts. We do not represent any undertaking to keep recipients advised of all legal developments. Prior results do not guarantee a similar outcome. © 2018 WilmerHale. All rights reserved.
Regarding the cybersecurity risks facing the DIB, including its supply chain, and the US military platforms, systems, networks, and capabilities, as well as recommendations for mitigating those risks.

The Order also requires designated agencies to:

- Identify authorities and capabilities to support cybersecurity efforts of CI entities at greatest risk, and engage and solicit input from those entities to evaluate whether and how the identified authorities and capabilities might be employed to support cyber risk management efforts, and any obstacles for doing so; and
- Engage in an “open and transparent process,” with appropriate stakeholders, to identify and promote action to improve internet and communications systems’ resilience and collaboration to reduce threats from automated and distributed attacks, such as botnets.

Cybersecurity of Federal Networks

The majority of the Executive Order focuses on improving IT security at Executive Branch agencies. However, defense and civilian agency contractors may see downstream effects from the Order’s requirements for their agency customers.

Specifically, the Order:

- Requires agency heads to implement the NIST Cybersecurity Framework. Contractors that have not yet implemented the NIST Cybersecurity Framework may see such a requirement imposed by agencies working to implement the Framework pursuant to the Order.
- Declares that agency heads will be held accountable by the President for “implementing risk management measures commensurate with the risk and magnitude of the harm that would result from unauthorized access, use, disclosure, disruption, modification, or destruction of IT data,” and for “ensuring that cybersecurity risk management processes are aligned with strategic, operational, and budgetary planning processes” in accordance with the Federal Information Security Management Act.
- Directs agency heads to show procurement preference for shared IT services, to the extent permitted by law, including email, cloud, and cybersecurity services. Contractors for IT services may face a changing business landscape in light of this preference, such as increased focus on General Services Administration IT Schedule contracts.
The Executive Order (the Order).

Order § 2(e). The assessment may be classified in whole or in part.

Order § 2(e). The report may be classified in whole or in part.

The Order explicitly references the process described in Section 9 of President Obama's Cybersecurity Executive Order (Executive Order 13,636 (Feb. 12, 2013)), and would not create a new designation process.

Order § 2(b). The agencies must report their findings to the President within 180 days, and update their findings annually thereafter.

Order § 2(d). The agencies must provide a preliminary report on this effort within 240 days, to be followed by a final report within one year.

Order § 1(c)(ii).

Order § 1(c)(i).
CONGRESS SEEKS TO FORCE (AND TIE) PRESIDENT'S HAND ON SANCTIONS THROUGH PASSAGE OF SIGNIFICANT NEW LAW CODIFYING AND EXPANDING U.S. SANCTIONS ON RUSSIA, NORTH KOREA, AND IRAN

To Our Clients and Friends:

Last month, we wrote to you about a "Blockbuster Week in U.S. Sanctions" during which the Senate overwhelmingly approved of a bill that would expand sanctions imposed against Iran and Russia.[1] Since then, the Senate bill proceeded to the House, where it was held for a little over a month, reportedly due to a "blue slip" violation—that is, a violation of the Constitutional provision requiring revenue provisions to originate in the House. This past week, the House introduced and overwhelmingly approved a version of that bill, titled the "Countering America's Adversaries through Sanctions Act," by a vote of 419-3.[2] Two days later, on July 27, the Senate voted to pass the House version of the bill by a vote of 98-2. The Trump administration has not yet committed to signing the bill into law, expressing some concerns about provisions in the bill that purport to limit the President's ability to terminate or waive sanctions. But even if President Trump were to veto the bill, it would still likely be enacted, given the overwhelming congressional support for the bill.

During the month that the bill—now "H.R. 3364"—was held in the House, two developments led to significant changes in the bill. First, on July 4, 2017, North Korea for the first time successfully tested an intercontinental ballistic missile. On July 14, House Majority Leader Kevin McCarthy announced his intention to add a North Korea sanctions bill, which had been previously passed by the House,[3] to the Iran-Russia sanctions bill. H.R. 3364 ultimately included a slightly modified version of the North Korea sanctions bill that had previously passed the House.

The other major development was that, during the month the sanctions bill was held in the House, U.S. and European business and policy leaders expressed concerns about certain provisions in the Russia-related portion of the bill. Leaders from the U.S. business and foreign policy communities, particularly from the U.S. energy industry, expressed concerns that certain provisions in the bill could have an adverse impact on U.S. businesses and workers—perhaps, ironically, to the benefit of Russian business interests. Leaders from the European business and foreign policy communities expressed concerns that the bill could weaken the business and diplomatic ties between the United States and Europe—again, perhaps, to the benefit of Russia. In response to these concerns, several small but significant changes were made to the bill.

This client alert details those changes and how they fit within the context of the bill's likely overall effects on the Iranian, Russian, and North Korean sanctions programs.
Iran

Title I of H.R. 3364, titled the "Countering Iran's Destabilizing Activities Act of 2017" ("CIDA"), imposes an additional set of significant sanctions against Iran. CIDA targets Iran's ballistic missile program, the Iranian Revolutionary Guard Corps, Iranian human rights abuses, and weapons transfers benefitting Iran. The bill also codifies the designations of persons pursuant to two executive orders and purports to limit the President's ability to remove those persons from the Specially Designated Nationals ("SDN") list. Although the Trump administration continues to criticize the Joint Comprehensive Plan of Action ("JCPOA") negotiated by the United States, the other permanent members of the U.N. Security Council, Germany and the European Union, these new sanctions do not undo or violate the JCPOA.

The provisions in the version of CIDA passed by Congress remain substantially identical to the Iran-related provisions in the bill as originally passed by the Senate. For a more detailed look at these provisions, consult our June 19 client alert, A Blockbuster Week in U.S. Sanctions.

Russia

Title II of H.R. 3364, titled the "Countering Russian Influence in Europe and Eurasia Act of 2017" ("CRIEEA"), imposes a significant set of sanctions against the Russian Federation. These sanctions are likely to have a wide-ranging impact on persons worldwide doing business with entities in the targeted sectors of the Russian economy.

The sanctions to be imposed under CRIEEA would be a dramatic expansion of what is already a complex sanctions and export control regime. CRIEEA would codify President Obama's executive orders imposing sanctions on Russia and also imposes new sanctions on additional activities and sectors of the Russian economy. The bill also imposes new restrictions on the President's ability to terminate or waive the sanctions that President Obama imposed on Russia and any new sanctions President Trump imposes on his own volition or as required by CRIEEA.

After the Senate passed the initial version of the sanctions bill, leaders from the U.S. and European business and foreign policy communities expressed concerns that the sanctions would, in several ways, have unintended adverse consequences on U.S. economic and diplomatic policy interests. In response to these concerns, several significant changes were made to the bill, as detailed below.

Codification and Expansion of Sectoral Sanctions

As noted in our June client alert, the sanctions bill expands the targets of the sectoral sanctions, initially set forth in President Obama's Executive Order 13662. Rather than "blacklisting" designated targets, these sanctions prohibit persons from engaging in certain activities with designated persons within certain sectors of the Russian economy—namely, the finance, energy, and defense sectors.

CRIEEA would expand the targeted sectors of the Russian economy to include state-owned entities in the railway, and metals and mining sectors. The initial Senate bill would also have added state-owned
entities in the shipping sector to the targeted sectors,[8] but the shipping sector was not included in the House version of the bill that was passed by Congress.

CRIEEA would alter Directives 1 and 2 of the sectoral sanctions, which prohibit U.S. persons from issuing certain types of new debt to designated Russian entities. Under the existing Directive 1, U.S. persons are prohibited from financing new debt with a maturity longer than 30 days to designated entities in Russia's financial services sector; CRIEEA would alter Directive 1 by reducing the maximum permitted maturity on new debt from 30 days to 14 days.[9] Under the existing Directive 2, U.S. persons are prohibited from financing new debt with a maturity longer than 90 days to designated entities in Russia's energy sector; CRIEEA would alter Directive 2 by reducing the maximum permitted maturity on new debt from 90 days to 60 days.[10] (The initial Senate bill would have reduced the maturity to 30 days, but this was revised to 60 in the House bill.)[11] These sanctions will increase the difficulty that those designated in the financial and energy sectors already face in financing and using debt in their day-to-day business operations.

The more significant changes to the Russian sectoral sanctions program are found in CRIEEA's expansion of Directive 4. The existing Directive 4 prohibits the provision of goods, support, or technology to designated Russian entities relating to the exploration or production for deepwater, Arctic offshore, or shale projects that have the potential to produce oil in the Russian Federation.[12] The expanded Directive 4 under CRIEEA removes this geographic limitation, and instead targets the specified types of oil projects worldwide in which a designated Russian person has a "controlling interest or a substantial non-controlling ownership interest in such a project defined as not less than a 33 percent interest."[13]

The language specifying the threshold at which the involvement of a designated person triggers Directive 4 was added to the House version of the bill. The initial version of CRIEEA would have prohibited U.S. persons from supporting such projects if a designated Russian entity had any involvement in the project, regardless of whether the project was capable of producing oil in the Russian Federation. This version of the bill prompted much concern in the U.S. energy sector, as it would have prohibited U.S. persons from taking part in any deepwater, Arctic offshore, or shale project, worldwide, in which a designated Russian firm had even a minimal amount of involvement. Ironically, the initial Senate bill could have allowed Russia to "weaponize" the sanctions to drive out U.S. interests from oil exploration and production projects.

The revised bill tempers this possibility by imposing a threshold below which involvement of a Russian entity does not trigger the prohibitions of Directive 4. Nevertheless, some uncertainty remains in the version of Directive 4 passed by Congress, particularly with respect to the definition of "controlling interest." Seeking to provide some clarity, Representatives Pete Sessions and Ed Royce engaged in a colloquy on the floor of the House, specifying their understanding that "controlling" would mean "the power to direct, determine, or resolve fundamental, operational, and financial decisions of an oil project through the ownership of a majority of the voting interests of the oil project."[14]

A second notable revision to CRIEEA's expanded Directive 4 is the addition of the qualifier that the expanded Directive 4 will only apply to new deepwater, Arctic offshore, and shale projects.[15] This
qualification clarifies that the Directive will not require U.S. energy businesses already engaged in projects that could be covered under the expanded Directive 4 to divest from such projects.

The version of CRIEEA passed by Congress also clarifies that the expanded Directives 1 and 2 will take effect 60 days after the enactment of the bill into law. The expanded Directive 4 will take effect 90 days after the enactment of the bill into law.

**Imposition of Secondary Sanctions**

As noted in our June client alert discussing the initial Senate version of CRIEEA, the willingness of the United States to impose so-called "secondary sanctions" on foreign persons and companies for conduct that may have no jurisdictional connection to the United States is a historical point of contention between the United States and its allies. Foreign persons that engage in prohibited activities under secondary sanctions can be effectively cut off from access to the goods, services, and financing of the U.S. market.

CRIEEA authorizes or mandates secondary sanctions with respect to several activities. The secondary sanctions provision in CRIEEA that received the most attention in the U.S. and European business communities was Section 232, which would authorize, but not require, the President to impose a set of secondary sanctions on any person providing a certain level of support to Russian energy export pipeline projects. European business and policy leaders objected to the bill's targeting of energy projects such as the Nord Stream 2 pipeline, in which several large European companies are involved and which is viewed by many European leaders as an essential part of Europe's long-term energy future. U.S. businesses feared that Section 232 could result in the imposition of sanctions on the subsidiaries of foreign companies that invest in Russian energy export pipelines, such as the Nord Stream 2 project.

These concerns appear to have affected change in the version of the bill that was passed by Congress, under which the President is only authorized to impose sanctions under Section 232 "in coordination with allies of the United States." Notably, Congress also revised an earlier portion of the bill to note that it is the "sense of Congress" that the President "should continue to uphold and seek unity with European and other key partners on sanctions implemented against the Russian Federation." These additions seemingly recognize the important role that the European Union plays in enforcing and imposing sanctions on Russia—indeed, it was in cooperation with the European Union that the sectoral sanctions imposed against Russia were implemented. Given the potential adverse effects the sanctions under Section 232 could have on the European Union, the requirement that the President only impose such sanctions "in coordination with allies of the United States" substantially reduces the likelihood that secondary sanctions under Section 232 will be imposed. In addition to this change, a small fix to Section 232 clarified that only Russian energy export pipeline projects—not all Russian energy pipeline projects—would be targeted if Section 232 were implemented.

The bill passed by Congress, like the initial Senate bill, requires the President to impose secondary sanctions on persons that engage in significant transactions with persons that are part of, or operate for or on behalf of, the defense or intelligence sectors of the Russian Federation. In a slight change from the initial Senate bill, the bill as passed allows the President to delay the imposition of these sanctions against a person that is reducing the number of transactions covered under these sanctions. Moreover,
the bill as passed requires the President to issue regulations further specifying which persons "are part of, or operate for or on behalf of, the defense and intelligence sectors of the Russian Federation."[22]

Several other provisions requiring the imposition of secondary changes in the bill as passed remain substantially unchanged from the initial Senate Bill. CRIEEA requires the President to impose sanctions on persons materially supporting certain efforts to undermine cybersecurity, persons investing in or otherwise facilitating the privatization of Russia's state-owned assets, persons providing certain types of support to the government of Syria, and persons evading sanctions.[23]

The menu of 12 possible secondary sanctions that can be imposed on persons violating the secondary sanctions provisions also remains substantially unchanged from the initial Senate version.[24] For an in-depth look at these possible sanctions, consult our client alert, *A Blockbuster Week in U.S. Sanctions.*

**Other Sanctions**

Other sanctions measures in CRIEEA were substantially unchanged between the initial and final versions of the bill. The bill as passed, like the initial Senate bill, requires the President to impose sanctions on persons engaged in, materially supporting, or providing financial services in support of "significant activities undermining cybersecurity of any person...on behalf of the Russian Federation."[25] Notably, on the same day that the Senate passed the initial version of CRIEEA, the EU announced it would be prepared to impose sanctions against "state and non-state actors" as part of a broader policy of combating malicious cyber activity.[26]

CRIEEA also amends the Ukraine Freedom Support Act ("UFSA") by making it mandatory unless the President determines that it is not in the national interest of the United States that the President impose sanctions on any person that "knowingly makes a significant investment" in a "special Russian crude oil project," which is defined as "a project intended to extract crude oil from—(A) the exclusive economic zone of the Russian Federation in waters more than 500 feet deep; (B) Russian Arctic offshore locations; or (C) shale formations located in the Russian Federation."[27] CRIEEA similarly modifies other parts of the UFSA to make additional sanctions mandatory rather than permissive.[28]

**Limitations on President's Ability to Terminate or Waive Sanctions**

One of the most unusual aspects of CRIEEA is the extent to which it seeks to limit the President's discretion to waive or terminate the sanctions or issue licenses that would "significantly alter[] United States' foreign policy with regard to the Russian Federation."[29] Under CRIEEA, before undertaking these actions, the President must first submit a report justifying the action to Congress, which then normally has 30 days to review the report.[30] Congress can pass a "joint resolution of disapproval" during this time, which, once passed by Congress, restricts the President from taking the described action for a period of 12 days.[31] If the President then vetoes the joint resolution of disapproval, he is still restricted from taking the action for an additional 10 days.[32] If the joint resolution of disapproval is ultimately enacted (presumably through a vote overriding the President's veto), the President cannot take the action.[33]
The constitutionality of these provisions is uncertain, given the restrictions on the President's discretion and, perhaps more problematically, the potential distortion of the constitutional processes for the enactment of legislation. Nevertheless, these provisions signal Congress's intention to take an active role in ensuring the vigorous enforcement of sanctions against Russia.

One significant change between the initial version of the bill and the bill as passed concerns the President's ability to issue licenses under the Russian sanctions. The bill requires the President to submit a report for congressional review whenever he engages in a "licensing action that significantly alters United States' foreign policy with regard to the Russian Federation."[34] Notably, however, the bill as passed by Congress contains a "Rule of Construction" clarifying that the bill "shall not be construed to require the submission of a report . . . with respect to the routine issuance of a license that does not significantly alter United States foreign policy with regard to the Russian Federation."[35] This is a significant addition to the bill, in that it leaves the President with substantial latitude to issue licenses authorizing activities that would otherwise result in sanctions being imposed. For instance, one concern raised by U.S. companies was that U.S. subsidiaries could be subject to sanctions under the bill's secondary sanctions provisions—not because of any activities the subsidiaries were themselves involved in, but rather because of the activities of their foreign parent companies. Under the Office of Foreign Asset Control's 50% rule, any entity that is more than 50% owned by a sanctioned entity is also treated as a sanctioned entity, and thus the activities of a foreign parent company could, in theory, result in a U.S. subsidiary being sanctioned, even if the U.S. subsidiary engaged in no sanctionable activities. The revision to the bill ensuring that the President retains his ability to issue routine licenses leaves open the possibility that the President could issue a license protecting the U.S. subsidiaries of sanctioned entities from being sanctioned under the 50% rule.

North Korea

On July 4, 2017, North Korea tested an intercontinental ballistic missile for the first time. Following this missile test, additional sanctions against North Korea were added to the sanctions bill targeting Iran and Russia. These additional sanctions are located in Title III of the bill, titled the "Korean Interdiction and Modernization of Sanctions Act" ("KIMSA").[36]

Under KIMSA, which is, in essence, a modified form of a bill passed by the House last March, the list of sanctionable activities with respect to North Korea is expanded. Under KIMSA, the President must designate persons that engage in:

- purchasing or acquiring any of various precious metals from North Korea;[37]
- selling or transferring rocket, aviation or jet fuel to North Korea;[38]
- providing fuel or supplies for designated North Korean vessels or aircrafts;[39]
- providing insurance services for vessels owned or controlled by the North Korean regime;[40] and
- maintaining a correspondent account with any North Korean financial institution.[41]
KIMSA also expands the list of activities triggering discretionary designations. Under KIMSA, the President may designate persons that engage in:

- purchasing or acquiring significant quantities of coal, iron, or iron ore from the government of North Korea;[42]
- purchasing or acquiring significant amounts of textiles from the government of North Korea;[43]
- facilitating a transfer of funds or property of the government of North Korea that materially contributes to any violation of an applicable United Nations Security Council Resolution;[44]
- transferring a significant amount of bulk cash, precious metals, gemstones, or other stores of value to or from the government of North Korea;[45]
- selling, transferring, or providing significant amounts of oil or natural gas to the government of North Korea;[46]
- engaging in or facilitating the online commercial activities of North Korea, including online gambling;[47]
- purchasing or acquiring fishing rights from the government of North Korea;[48]
- purchasing or acquiring significant types or amounts of food or agricultural products from the government of North Korea;[49]
- engaging in, facilitating, or being responsible for the exportation of workers from North Korea in a manner intended to generate significant revenue for the government or ruling party of North Korea;[50]
- conducting a significant transaction in North Korea's transportation, mining, energy, or financial services industries;[51] and
- facilitating the operation of any branch, subsidiary, or office of a North Korean financial institution.[52]

The bill also takes steps to strengthen measures aimed at North Korean correspondent accounts,[53] the North Korean shipping sector,[54] and North Korea's use of forced labor.[55] Finally, the bill requires the Secretary of State to submit a determination as to whether North Korea meets the criteria for designation as a state sponsor of terrorism.[56]

These sanctions are a signal that Congress is continuing to search for ways to isolate the North Korean regime and limit its ability to finance its operations. Given the broad range of criteria these sanctions now introduce for mandatory or discretionary sanctions designations, KIMSA will enable the President to impose sanctions on many Chinese companies and persons that have historically provided the specified kinds of support to North Korea. We expect the Chinese government to have a strong reaction.
to this expansion of North Korean sanctions either upon enactment of the bill or once the President begins to designate Chinese entities under the law.

Conclusion

The overwhelming congressional support for expanded sanctions on Iran, Russia, and North Korea underscores the extent to which there exists a political appetite for the extensive use and vigorous enforcement of sanctions as a tool of foreign policy. It also signals Congress's appetite for taking a direct role in the imposition of sanctions, potentially setting up a conflict with the executive branch over the proper scope of Congress's role in foreign policy.

Historically, the U.S. sanctions regime has been shaped by the competing demands of the Executive and Legislative branches. Once codified, it has been tremendously difficult to muster congressional support for further modification of the law. For this reason, the sanctions relief offered by the JCPOA takes the form of a complicated licensing scheme designed to permit activities otherwise prohibited by various statutes. In passing the new Iran, Russia and North Korea sanctions, Congress has made it significantly more difficult to modify U.S. sanctions policy, setting the stage for a similarly arduous process of retooling policies if and when these regimes change.

As the sanctions legislation makes it way towards the President's desk and potential enactment into law, risks abound for entities and individuals with cross-border exposure, given the rapidly changing nature of sanctions policies and priorities.

The Gibson Dunn team will continue to provide frequent updates on this and any other sanctions programs that may receive alterations or enhancements in the time ahead.


[7] Id. § 223(a).


[10] Id. § 223(c).


[16] See id. § 223(b)-(c).

[17] See id. § 223(d).

[18] See id. § 232.

[19] Id.

[20] See id. § 212.

[21] Id. § 231.

[22] See id. § 231(c)-(d).

[23] See id. §§ 224(a)(2) (cybersecurity), 228 (evasion), 233(a) (state-owned assets), 234 (Syria).


[25] Id. § 224(a).


[27] H.R. 3364, § 225.

[28] See id. § 226.

[29] Id. § 216(a)(2)(A)(iii).
Id. § 216(b)(1). Congress has 60 days if the report is submitted between July 10 and September 7.

Id. § 216(b)(4).

Id. § 216(b)(5).

Id. § 216(b)(6).

Id. § 216(b)(1).

Id. § 216(a)(6).

Id. § 301.

Id. § 311(a)(10).

Id. § 311(a)(11).

Id. § 311(a)(12).

Id. § 311(a)(13).

Id. § 311(a)(14).

Id. § 311(b)(1)(D).

Id. § 311(b)(1)(E).

Id. § 311(b)(1)(F).

Id. § 311(b)(1)(G).

Id. § 311(b)(1)(H).

Id. § 311(b)(1)(I).

Id. § 311(b)(1)(J).

Id. § 311(b)(1)(K).

Id. § 311(b)(1)(L).

Id. § 311(b)(1)(M).

Id. § 311(b)(1)(N).
The following Gibson Dunn lawyers assisted in preparing this client update: Judith Alison Lee, Adam Smith, Henry Phillips, Stephanie Connor, and Christopher Timura.

Gibson Dunn’s lawyers are available to assist in addressing any questions you may have regarding the above developments. Please contact the Gibson Dunn lawyer with whom you usually work, the authors, or any of the following leaders and members of the firm’s International Trade Group:

**United States:**

- Judith A. Lee - Co-Chair, International Trade Practice, Washington, D.C. (+1 202-887-3591, jalee@gibsondunn.com)
- Ronald Kirk - Co-Chair, International Trade Practice, Dallas (+1 214-698-3295, rkirk@gibsondunn.com)
- Caroline Krass - Chair, National Security Practice, Washington, D.C. (+1 202-887-3784, ckrass@gibsondunn.com)
- Alexander H. Southwell - Chair, PCCP Practice, New York (+1 212-351-3901, asouthwell@gibsondunn.com)
- Jose W. Fernandez - New York (+1 212-351-2376, jfernandez@gibsondunn.com)
- Marcellus A. McRae - Los Angeles (+1 213-229-7675, mmcrae@gibsondunn.com)
- Daniel P. Chung - Washington, D.C. (+1 202-887-3729, dchung@gibsondunn.com)
- Adam M. Smith - Washington, D.C. (+1 202-887-3547, asmith@gibsondunn.com)
- Christopher T. Timura - Washington, D.C. (+1 202-887-3690, ctimura@gibsondunn.com)
- Stephanie L. Connor - Washington, D.C. (+1 202-955-8586, sconnor@gibsondunn.com)
- Kamola Kobildjanova - Palo Alto (+1 650-849-5291, kkobildjanova@gibsondunn.com)

**Europe:**

- Peter Alexiadis - Brussels (+32 2 554 72 00, palexiadis@gibsondunn.com)
- Attila Borsos - Brussels (+32 2 554 72 10, aborsos@gibsondunn.com)
- Patrick Doris - London (+44 (0)207 071 4276, pdoris@gibsondunn.com)
- Penny Madden - London (+44 (0)207 071 4226, pmadden@gibsondunn.com)
- Benno Schwarz - Munich (+49 89 189 33 110, bschwarz@gibsondunn.com)
- Mark Handley - London (+44 (0)207 071 4277, mhandley@gibsondunn.com)

© 2017 Gibson, Dunn & Crutcher LLP

Attorney Advertising: The enclosed materials have been prepared for general informational purposes only and are not intended as legal advice.
Gibson, Dunn & Crutcher offers clients worldwide a wide array of the most current, experienced and connected practitioners in national security to navigate the physical, cyber, and financial threats global businesses face, as well as the government efforts to address them.

Our team includes an unparalleled collection of former U.S. federal officials who served in senior capacities at the Central Intelligence Agency (CIA), National Security Council (NSC), White House, U.S. Department of Justice (DOJ), Department of the Treasury (Treasury), Department of State, Office of Management and Budget (OMB), and other agencies that make, implement, and enforce U.S. security laws and regulations.

Gibson Dunn’s National Security Practice Group is uniquely equipped to address every aspect of a national security legal or public policy issue. We offer a comprehensive range of services, including a counseling practice focused on businesses operating internationally as they work to comply with federal laws and regulations governing cross-border economic transactions. We represent government contractors and other companies involved in classified government programs. As the U.S. government increasingly looks to the private sector to assist with counterterrorism and other national security efforts, including by making demands for information and imposing security mandates, our lawyers have been at the forefront of cutting-edge questions concerning government access to data. And we handle the special challenges presented by civil litigation as well as criminal, regulatory, and congressional investigations that raise national security issues.

National security matters are among the most important and difficult legal issues confronting all major corporations. Our team understands how the U.S. government approaches the full range of national security legal questions – including military and intelligence matters. We can help clients navigate this changing and challenging landscape.

Our Team

Our National Security Practice Group draws on the experience and knowledge of many of the firm’s world-renowned practice groups, including Crisis Management, Financial Institutions, Government Contracts, International Trade, Labor and Employment, Public Policy, Litigation, and White Collar Defense and Investigations. The national security team includes partners, senior lawyers and others with extensive government experience, top secret security clearances and wide-ranging capabilities in the full spectrum of national security issues.
As noted, many of our team members have served in senior U.S. government national security positions, including high-level positions in the DOJ and Treasury, the CIA, the White House, NSC, and OMB. Among many others, we have a former General Counsel of the CIA, a former Acting Associate Attorney General of the United States (the third-ranking position at DOJ), a former Solicitor General of the United States, three former U.S. Attorneys and many other former federal prosecutors, a former legal advisor to the NSC, a former Special Counsel to the President for National Security Affairs, a former Associate Director of OMB, and a former Assistant Secretary of State for Economic, Energy and Business Affairs. While in government, our lawyers worked closely with every national security-related agency, including the Office of the Director of National Intelligence, the CIA, the Department of Defense (DoD), the National Security Agency (NSA), the Defense Security Service, the Federal Bureau of Investigation and the Department of Homeland Security. Gibson Dunn is also fortunate, and proud, to have on our team over a dozen former military officers.

A number of Gibson Dunn team members either currently hold, or have recently held (and can reactivate on short notice), security clearances through the Top Secret/SCI level. This allows us to get involved rapidly in matters where access to classified material is required and time is of the essence.

We advise clients in the areas of:

- Classified programs
- Government access to information
- National security litigation and investigations
- Cross-border economic activities
- Cybersecurity
- National security public policy

**Classified Programs**

When corporate executives are approached by the U.S. intelligence community (including the CIA and the NSA), or the DoD, and asked to assist with a sensitive and classified project in the interest of national security, they look for legal guidance about how to protect the corporation against legal risk and appropriately manage the project internally. Gibson Dunn has the experience to provide that kind of targeted legal advice concerning highly sensitive relationships with the intelligence community and other U.S. national security agencies.

The lawyers in our government contracts and aerospace and related technologies groups have decades of experience providing counseling and litigation services to defense industry and commercial clients regarding national security implications of their business relationships with the U.S. government. Several of our lawyers handled contracting issues or litigation concerning classified programs while in government, so we can bring special insight in addressing these issues. We have been involved in numerous hearings and disputes, conducted entirely in the “black world,” on precedent-setting cases involving the State Secrets privilege and the Defense Production Act. We have represented clients in matters involving the CIA, the NSA, the National Reconnaissance Office, the National Geospatial-Intelligence Agency, and other federal agencies involved in intelligence gathering and analysis. We regularly work with companies before the Defense Security Service and assist them in satisfying the regulatory requirements of the National Industrial Security Program (NISP) that are driven by foreign investments or acquisitions. We can advise
clients on how best to address any challenge in protecting classified information associated with classified contracts. We are also experienced in advising clients operating in areas of conflict. Gibson Dunn lawyers have represented clients across the entire spectrum of battlefield conditions, from peace operations and support and stability operations to full-scale armed conflict and post-conflict environments in locations such as Afghanistan, Iraq and Kuwait. Our lawyers have extensive experience representing contractors and individuals in theater and are highly knowledgeable about the laws, regulations and challenges clients face when performing work in a battlefield environment.

Government Access to Information

Government demands for information from companies through voluntary requests, warrants, or other compulsory process increasingly present novel questions given technological advances in encryption, the global nature of data storage, and government surveillance capabilities and authorities.

Companies have a strong interest in protecting the privacy of their customer information, and encryption is of course a vital tool for consumers and thus for the commercial interests of providers. On the other hand, law enforcement and intelligence agencies have a legitimate need for assistance in urgent national security investigations and in obtaining compliance with valid court orders for personal information. U.S. companies are subject to surveillance orders from law enforcement and the intelligence community under a number of statutes, including the Wiretap Act, the Foreign Intelligence Surveillance Act (FISA), the Electronic Communications Privacy Act (ECPA) and the Stored Communications Act (SCA). In addition, the Wiretap Act and SCA allow civil lawsuits challenging the use of personal information brought by individuals or through class actions.

U.S. officials in recent years have expressed the view that cooperation from the private sector will be increasingly necessary for the government to obtain the information it needs to pursue urgent national security investigations and other law enforcement operations. In the absence of legislation to strike the right balance in the context of current technology, these questions increasingly will be resolved through negotiation with the Executive Branch or, failing that, in the courts. Gibson Dunn has played a leading role in cases addressing these questions. We can challenge surveillance orders, including requirements that customers not be notified of such orders in certain circumstances. With our unparalleled combination of experience in representing tech companies and managing these issues in government, our national security practice group tackles and solves the most difficult data access questions our clients face.

National Security Litigation and Investigations

Increasingly, our clients are confronting civil litigation and criminal investigations that raise national security issues, and Gibson Dunn has deep experience in such matters. With our Appellate and Constitutional Law Practice Group, our team is prepared to handle difficult legal questions at every level. And the Transnational Litigation Practice Group specializes in protecting clients against claims in U.S. and other courts stemming from overseas activities, as well as reducing and eliminating the risks posed by foreign litigation.
National security litigation and investigation matters often turn on classified information, and our lawyers are well versed in the complex procedures for handling classified information in the federal courts, including the Classified Information Procedures Act and the State Secrets privilege. We are also experienced in specialized doctrines that may arise in such cases, such as sovereign and official immunity, and executive and deliberative process privilege. Our lawyers have experience with litigation involving misuse of classified information, including potential compulsion of reporters and criminal prosecutions, as well as security clearance disputes.

National security litigation and investigations often present situations in which Gibson Dunn’s experience in crisis management is critical, including executing a strategic communication plan, responding to potential congressional interest, engaging with the Executive Branch, and otherwise guiding clients through difficult events.

Cross-Border Economic Activities

Our national security team helps global businesses comply with federal laws and regulations governing cross-border economic activities.

Foreign Investment. The continuing rise in cross-border transactions, in combination with the U.S. government’s attention to national security considerations associated with such transactions, require clients increasingly to turn to counsel for advice in complying with federal regulations designed to protect national security. For example, review by the Committee on Foreign Investment in the United States (CFIUS) will continue to be a critical focus of the U.S. government as national security concerns escalate about investment by certain countries and as foreign investment affects critical U.S. technologies and infrastructure. Gibson Dunn has extensive experience assisting both U.S. and foreign companies in transactions subject to CFIUS review. In addition, when companies are determined to be under foreign ownership, control or influence (FOCI), we have assisted in obtaining security clearances for company employees and handling other issues to enable our clients to obtain or retain federal government work.

Economic Sanctions and Embargoes (Office of Foreign Assets Control). Gibson Dunn advises both U.S. and foreign clients engaged in international commerce on a wide array of issues stemming from compliance with U.S. sanctions and embargoes administered and enforced by Treasury, Office of Foreign Assets Control (OFAC), and the Department of State. Our lawyers have substantial experience assisting companies with conducting compliance assessments and risk audits, advising on the scope and applicability of particular sanctions regulations, liaising with key officials and regulators to obtain guidance or insight on policy focus and direction, and preparing requests for official advisory opinions or licenses. When a company is concerned that sanctions violations may have occurred, we frequently assist in the internal investigation to uncover the facts, prepare and submit any self-disclosures and, if necessary, represent the company during the course of a subsequent government investigation and/or enforcement action.

Export Controls. Our international trade and government contracts groups advise clients on a wide range of U.S. government restrictions on the export of goods and technologies. Our clients regularly consult us for advice on complying with the Export Administration Regulations and the International Traffic in Arms Regulations. We assist clients in developing effective programs to ensure compliance with license conditions and reporting requirements, and in responding to enforcement actions. We also assist clients in
obtaining licenses and other authorizations from the U.S. Commerce Department’s Bureau of Industry and Security (BIS) and the State Department’s Directorate of Defense Trade Controls (DDTC).

**Foreign Corrupt Practices Act and International Bribery Compliance and Investigations.** In recent years, the DOJ has identified FCPA enforcement as a national security issue tied to the global fight against terrorism. As clients expand business opportunities overseas, the risk of corruption and bribery has increased exponentially. Gibson Dunn has represented clients in investigations involving alleged FCPA violations since the inception of the law. Our lawyers regularly advise clients on all aspects of the FCPA and multinational corruption issues, including: initial advice, training, establishment and implementation of effective compliance programs, representation in internal and government investigations involving suspected FCPA and international corruption violations, Securities and Exchange Commission implications of FCPA issues, merger and acquisition due diligence related to corporate transactions, and special engagements such as serving in corporate compliance monitorships.

**Anti-Money Laundering.** Members of our financial institutions and white collar defense groups have extensive knowledge of U.S. and foreign anti-money laundering laws and regulations, both as government lawyers and private practitioners. We provide advice on anti-money laundering and forfeiture laws, currency transaction reporting, the identification and reporting of suspicious activity, wire transfer and other record-keeping requirements. We advise and represent financial institutions in audits and exams conducted by a number of regulators and help prepare our clients to successfully complete the exercises in the most cost-effective manner. We also defend financial institutions and other regulated entities against investigations and prosecutions for failure to prevent or report illegal money laundering and failure to follow the “know-your-customer” requirements and other requirements of the USA PATRIOT ACT. Our clients include banks and financial service companies, as well as major international businesses, including manufacturers and distributors of appliances, consumer electronics, computers, and other goods.

**Cybersecurity**

The U.S. government’s focus on the risks of cyber attacks and the need for preparedness, together with the increased publicity surrounding such attacks, has resulted in a strong desire by clients to mitigate the associated legal risk. Gibson Dunn’s Privacy, Cybersecurity and Consumer Protection Practice Group has a demonstrated history of helping companies successfully navigate the complex and rapidly evolving laws, regulations, and industry best practices relating to privacy, cybersecurity and consumer protection. We have substantial experience assisting companies with all facets of cybersecurity, including helping clients develop policies and procedures before an incident arises, as well as counseling clients through the important steps that must occur immediately after breach situations and navigating the federal and state government investigations and private litigation that increasingly accompany cybersecurity incidents. Due diligence regarding cybersecurity considerations has become increasingly important in the context of mergers and acquisitions, and we have great familiarity with handling these issues. And for those clients that work with the U.S. government, we advise regarding compliance with rules related to safeguarding government information from cyber attacks. Gibson Dunn is proud to have been named one of the seven “Law Firms Best at Cybersecurity” in BTI Consulting Group’s 2017 report based on in-depth interviews with more than 320 corporate counsel at the world’s largest companies.

Our capabilities are global. Cybersecurity and privacy are global issues, and Gibson Dunn draws on its international team to seamlessly advise clients on sophisticated multijurisdictional matters. The practice
group includes lawyers in Brussels, London, Paris, Munich, Beijing, Singapore and Hong Kong who are not only exceptionally knowledgeable on relevant data protection and privacy laws at the national level, but are experienced in advising companies on European Union and Chinese developments, such as the EU Privacy Shield and upcoming General Data Protection Regulation, and coordinating multinational approaches.

National Security Public Policy

Members of our public policy group have extensive experience with effecting public policy change on national security matters. We are prepared to help build strategies regarding national security policy matters pending before both the U.S. Congress and the Executive Branch. Our team includes former government officials who were responsible for providing advice to White House and Executive Branch policymakers and addressing congressional oversight of national security programs. We can bring that perspective to bear in advising clients in their interactions with the federal government on policy and regulatory matters. We have prepared clients to testify before congressional committees, assisted foreign firms in obtaining approvals from regulatory agencies to acquire U.S. defense contractors, defended technology clients against hostile takeover bids based on national security concerns, and assisted defense and aerospace firms with advocacy issues both in the United States and in foreign countries.
Late last week, a group of technology industry leaders, researchers, and former policymakers comprising the President’s Council on Science and Technology (“PCAST”) issued a report entitled *Ensuring Long-Term U.S. Leadership in Semiconductors*. The report concludes that China is seeking to reshape the global semiconductor industry in a manner that threatens U.S. leadership in that sector. Given the criticality of the semiconductor industry to U.S. national security, the report recommends a three-part strategy to (1) counter Chinese policy actions, (2) improve the business environment for U.S. semiconductor companies, and (3) help “catalyze transformative semiconductor innovation over the next decade.”

The report does not reflect a new focus on the semiconductor industry; rather, it is the culmination of several years’ worth of study and analysis throughout the executive branch of a vexing industrial policy challenge—the apparent hegemony that China seeks to exert over the semiconductor industry, and the associated implications for U.S. national and economic security. Nor does the report itself mark a U.S. change in law or regulation, or even announce a specific policy shift. That said, the report is notable because it lays a marker for even greater scrutiny of Chinese investments in the semiconductor and adjacent sectors by the Committee on Foreign Investment in the United States (“CFIUS”), and conveys a message for the incoming Trump Administration policymakers who will inherit the challenges addressed in the report. In this regard, we expect there will be considerable discussion within the new Administration and the Congress of the report’s suggested legal and policy responses, including expanded multilateral collaboration on export controls and investment reviews, and the use of other trade tools to combat Chinese government-supported activity that the report characterizes as market-distorting. We discuss certain aspects of the report, and some takeaways, below.

**Report Conclusions and U.S. Government Concerns Regarding Industry Trends**

Among the report’s more notable conclusions are the following:

- China has made a concerted effort to re-shape the semiconductor industry through a range of policies backed by over $100 billion in government funds. These policies include (i) subsidies intended to encourage foreign companies to locate facilities in China and domestic firms to acquire foreign companies and technologies, and (ii) “zero-sum” tactics, such as requiring domestic customers to purchase from Chinese suppliers,
forcing technology transfers in exchange for market access, and misappropriating intellectual property.

- These Chinese policies are “distorting the market” and “put U.S. national security at risk.” Maintaining a leading position in semiconductors is critical not only to “defense systems and U.S. military strength,” but also to mitigating cybersecurity risks given the ubiquity of semiconductors in electronics.

- The U.S. government should “attempt to influence Chinese behavior” by improving transparency around Chinese policy, working with allies to “coordinate and strengthen inward investment security and export controls,” and “responding firmly and consistently to Chinese violations of international agreements.”

As we have reported previously, the industry trends that the report identifies, and the high-level focus on the issue from U.S. policymakers, have been long in the making. Commerce Department Secretary Penny Pritzker took the unusual step of making a major policy address late last year focused specifically on the semiconductor industry and China, noting the Chinese government’s announced objective of spending $150 billion to increase China’s self-supply of integrated circuits to 70 percent by 2025. For its part, CFIUS has for some time been conducting a thorough risk-based analysis for every transaction in the semiconductor sector involving China—and even some that do not involve China directly, but where there could be Chinese interests, such as where there are Chinese operations or customers of the existing U.S. business—regardless of how benign the transaction may appear.

We also have seen U.S. concerns regarding China’s activities in the semiconductor industry play out in regulatory actions, including through CFIUS. Most recently, on December 2, 2016, President Obama issued an Executive Order prohibiting the acquisition of the U.S. business of Aixtron SE (“Aixtron”), a maker of semiconductor manufacturing equipment, by a company ultimately owned by Chinese investors. The order marked the first time that a president formally has utilized the authority of the CFIUS statute to block a foreign acquisition prior to consummation of the transaction. In a statement, the Department of the Treasury—which chairs CFIUS—observed that “the national security risk posed by the transaction relates, among other things, to the military applications of the overall technical body of knowledge and experience of Aixtron, a producer and innovator of semiconductor manufacturing equipment and technology.”

While the industry trends identified by the report are not new, the report characterizes the U.S. national security concerns about those trends more starkly than we previously have seen conveyed publicly. We accordingly expect the report will intensify the focus within the U.S. government on addressing the perceived risks to the U.S. semiconductor industry presented by China’s actions.

---

1 For more information, please see our client alert on the Aixtron executive order here.
Lessons for Cross-Border Acquisitions and Investments

We see the following takeaways from the report for companies in the semiconductor industry:

1. **Semiconductor Transactions Involving China Will Receive Greater Scrutiny, Including from CFIUS.** The report notes that Chinese firms have been "increasingly active in the acquisition space" and that "the most likely avenue for growth will be acquisition of global players (or divisions of them) in the United States, Europe, or Japan." CFIUS will continue to scrutinize semiconductor transactions involving Chinese parties to determine whether the transactions present any risk to U.S. national security. Notably, the report recommends that the U.S. government continue to maintain its policy focus, including through CFIUS, exclusively on the national security impacts of China’s actions, and not utilize tools such as CFIUS to pursue economic—as opposed to national security—goals. At the same time, in addressing “Chinese policies that distort the global market by limiting access of U.S. companies and U.S. exports to China’s large and growing semiconductor market,” the report says:

   One way to respond would be to tie U.S. assessments of the national-security threats posed by particular technology exports, investments, and contracts to Chinese policy. (For example, if China pursues a policy of undermining cutting-edge, defense-critical U.S.-based companies by flooding markets using government support, that should alter U.S. assessments of whether Chinese acquisitions of the capabilities required to do so are acceptable.)

   The main goal here should be to deter dangerous Chinese actions; this means that the United States will need to be more open and clear about how its investment and export restrictions actually work. If, however, this effort to deter fails, changes in U.S. national security threat assessments will presumably lead to changes in the specific exports and acquisitions allowed by the U.S. government.

   This, to our knowledge, is the first time that a report sponsored by an executive branch agency has overtly recommended the inclusion of a reciprocity-like factor in a national security review. While the recommendation does not reflect official policy, it does reflect views being expressed within industry and government discussions, to which government officials in the CFIUS process are sure to be attuned.

2. **Securing Approval for Transactions with Chinese Parties in the Semiconductor Industry Will Be Harder—but Not Impossible—Absent a Change in Policy by the Trump Administration.** As a result of increased scrutiny and sensitivity, securing CFIUS approval for semiconductor transactions will likely be even more challenging than it already has been in recent years. Transaction parties should be prepared to respond to extensive inquiries relating to the transaction and to work aggressively with counsel to identify and address any potential national security issues arising from the transaction, ideally before the transaction is filed with CFIUS. That does not mean, however, that semiconductor transactions with Chinese parties will be impossible. CFIUS continues to approve such transactions where there are no perceived national security risks, or where it determines that the risks can be fully mitigated—and, as a legal matter, CFIUS must continue to comply with its clear statutory mandate only to take action where transactions threaten to impair U.S. national security.
3. **The Most Intense Focus Will Be on Leading Edge Technologies.** The report notes that leading edge semiconductor technology is “critical to sustaining a national-security edge,” and concludes that policymakers should focus principally on such technologies. As a general rule, therefore, we expect transactions involving leading-edge technology to receive more scrutiny than those involving less advanced or more commoditized technologies. In transactions involving less advanced technology, CFIUS is more likely to focus on other issues such as supply security and integrity.

4. **The U.S. Government Is Likely to Use Available Authorities, Including CFIUS, to Address Perceived Risks from Transactions, Including Transactions That Are Principally Outside the United States.** Given the globalized nature of the semiconductor industry, even transactions with only a tangential connection to the United States may be viewed as presenting risks to U.S. national security. The Aixtron transaction, for example, principally involved assets that were outside the United States, and while the President’s order carefully focused on the U.S. business of Aixtron, its scope potentially could encompass assets of Aixtron outside the United States, to the extent those assets are “used in, or owned for the use in or benefit of” Aixtron, Inc.’s activities in the United States. As a result, transaction parties should not assume that they do not need to file with CFIUS because a business has only a tangential connection to the United States, but rather confer with experienced CFIUS counsel early in the process to identify potential issues and solutions.

5. **The U.S. Government Is Likely to Work More Closely with Allied Governments to Address Perceived National Security Risks Arising in the Semiconductor Industry, Including on Export Controls.** The report recommends that the United States work with allies to strengthen global export control and inward investment security. In the Aixtron case, media reports suggested that the German government withdrew its approval of the transaction after receiving previously-unknown security-related information from U.S. intelligence officials. We expect that U.S. authorities will increasingly work with counterparts in U.S.-ally countries that are leaders in semiconductor technology, including Japan, Korea, Germany, and the Netherlands, among others. In the export controls area, this could take the form of seeking agreement with such countries on approaches to licensing the export of technology to China or Chinese persons, and pressing for greater enforcement of export control rules when there are violations involving China or Chinese persons.

6. **There Is Increased Potential for Legislative Activity in the U.S.** While the PCAST report generally seems supportive of decreasing regulatory burdens on the semiconductor industry—and it does not call for any reform of CFIUS—the concerns raised by the report could be cited by those exploring CFIUS-related legislation as further grounds for reforms targeted at China. As we have written previously, while the legislative landscape in the new Congress is still being defined, there is a greater chance now than at any time in the last decade of potential legislation to amend CFIUS. Among the reforms being considered include amendments that would expand CFIUS to cover licensing transactions or transactions that result in a contribution of technology or know-how from the U.S. to overseas joint ventures—which are being driven in part by concerns over transfers of semiconductor expertise to China. Thus, while the report does not call for legislation, it could be cited to substantiate potential reforms that will be debated in the next several months.
We hope that you find this analysis useful. Please do not hesitate to contact the following members of our International Trade Practice Group if you would like to discuss any aspect of the foregoing in further detail:

<table>
<thead>
<tr>
<th>Name</th>
<th>Phone Number</th>
<th>Email Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark Plotkin</td>
<td>+1 202 662 5656</td>
<td><a href="mailto:mplotkin@cov.com">mplotkin@cov.com</a></td>
</tr>
<tr>
<td>David Fagan</td>
<td>+1 202 662 5291</td>
<td><a href="mailto:dfagan@cov.com">dfagan@cov.com</a></td>
</tr>
<tr>
<td>Stuart Eizenstat</td>
<td>+1 202 662 5519</td>
<td><a href="mailto:seizenstat@cov.com">seizenstat@cov.com</a></td>
</tr>
<tr>
<td>Alan Larson</td>
<td>+1 202 662 5756</td>
<td><a href="mailto:alarson@cov.com">alarson@cov.com</a></td>
</tr>
<tr>
<td>Peter Lichtenbaum</td>
<td>+1 202 662 5557</td>
<td><a href="mailto:plichtenbaum@cov.com">plichtenbaum@cov.com</a></td>
</tr>
<tr>
<td>Mark Plotkin</td>
<td>+1 202 662 5656</td>
<td><a href="mailto:mplotkin@cov.com">mplotkin@cov.com</a></td>
</tr>
<tr>
<td>David Fagan</td>
<td>+1 202 662 5291</td>
<td><a href="mailto:dfagan@cov.com">dfagan@cov.com</a></td>
</tr>
<tr>
<td>Stuart Eizenstat</td>
<td>+1 202 662 5519</td>
<td><a href="mailto:seizenstat@cov.com">seizenstat@cov.com</a></td>
</tr>
<tr>
<td>Alan Larson</td>
<td>+1 202 662 5756</td>
<td><a href="mailto:alarson@cov.com">alarson@cov.com</a></td>
</tr>
<tr>
<td>Peter Lichtenbaum</td>
<td>+1 202 662 5557</td>
<td><a href="mailto:plichtenbaum@cov.com">plichtenbaum@cov.com</a></td>
</tr>
<tr>
<td>John Veroneau</td>
<td>+1 202 662 5034</td>
<td><a href="mailto:jveroneau@cov.com">jveroneau@cov.com</a></td>
</tr>
<tr>
<td>Roger Zakheim</td>
<td>+1 202 662 5959</td>
<td><a href="mailto:rzakheim@cov.com">rzakheim@cov.com</a></td>
</tr>
<tr>
<td>Tim Stratford</td>
<td>+86 10 5910 0508</td>
<td><a href="mailto:tstratford@cov.com">tstratford@cov.com</a></td>
</tr>
<tr>
<td>Gina Vetere</td>
<td>+1 202 662 5647</td>
<td><a href="mailto:gvetere@cov.com">gvetere@cov.com</a></td>
</tr>
<tr>
<td>Damara Chambers</td>
<td>+1 202 662 5279</td>
<td><a href="mailto:dchambers@cov.com">dchambers@cov.com</a></td>
</tr>
<tr>
<td>Heather Finstuen</td>
<td>+1 202 662 5823</td>
<td><a href="mailto:hfinstuen@cov.com">hfinstuen@cov.com</a></td>
</tr>
<tr>
<td>Meena Sharma</td>
<td>+1 202 662 5724</td>
<td><a href="mailto:msharma@cov.com">msharma@cov.com</a></td>
</tr>
<tr>
<td>Jonathan Wakely</td>
<td>+1 202 662 5387</td>
<td><a href="mailto:jwakely@cov.com">jwakely@cov.com</a></td>
</tr>
<tr>
<td>Ingrid Price</td>
<td>+1 202 662 5539</td>
<td><a href="mailto:jprice@cov.com">jprice@cov.com</a></td>
</tr>
<tr>
<td>Andrew Vaden</td>
<td>+1 202 662 5755</td>
<td><a href="mailto:avaden@cov.com">avaden@cov.com</a></td>
</tr>
</tbody>
</table>

This information is not intended as legal advice. Readers should seek specific legal advice before acting with regard to the subjects mentioned herein.

Covington & Burling LLP, an international law firm, provides corporate, litigation and regulatory expertise to enable clients to achieve their goals. This communication is intended to bring relevant developments to our clients and other interested colleagues. Please send an email to unsubscribe@cov.com if you do not wish to receive future emails or electronic alerts.
TRUMP ADMINISTRATION PRIORITIZATION OF CYBERSECURITY

To Our Clients and Friends:

Headlines regarding the "WannaCry" ransomware attack dominated recent global press. But a less-heralded cybersecurity development from the U.S. federal government—President Donald Trump's signing of an expansive Executive Order on May 11, 2017—is likely to prove far more significant in the long run. The Executive Order, titled "Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure," focuses on cybersecurity protections for federal agencies and critical infrastructure, but the implications of the Order will be significant for virtually all medium- and large-sized businesses operating in the United States, the latest in a string of notable developments at the federal, state, and local level regarding cybersecurity.[1]

Setting the Stage

Following the first high-profile mention of cybersecurity in a U.S. State of the Union address, in February 2013, then-President Barack Obama issued an Executive Order calling for the development of a voluntary framework to establish common standards for managing cybersecurity risk.[2] Pursuant to this Order, a year later, following extensive industry collaboration, the U.S. Department of Commerce's National Institute of Standards and Technology ("NIST") issued a framework for improving critical infrastructure cybersecurity in February 2014.[3] The NIST Framework created a set of industry standards and best practices, including a shared vocabulary about cybersecurity to help decision-makers manage cybersecurity using a risk-based approach.

Following up on this initiative, in February 2016 the Obama Administration unveiled a Cybersecurity National Action Plan ("CNAP") to put in place a long-term strategy to enhance cybersecurity awareness and protections. The CNAP included efforts to modernize the federal government's aging information technology infrastructure and the establishment of a Commission on Enhancing National Cybersecurity, which would involve leaders from the private sector in making recommendations regarding cybersecurity policy.[4] On December 1, 2016, the Commission on Enhancing National Cybersecurity issued recommendations in its Report on Securing and Growing the Digital Economy.[5]

Upon taking office, President Trump announced that his Administration would take steps to improve the United States' cybersecurity posture—an announcement that came after an election in which alleged cybersecurity failures, such as the disclosure of emails covertly collected from the Democratic National Committee, were prominently featured in news reports. Until the May 11 Executive Order was finalized, however, the specifics of the Trump Administration's plans remained unclear. Some observers pointed to the appointment of veteran national security advisers to key cybersecurity-related posts in the White
House and Department of Homeland Security as evidence that the Trump Administration likely would continue on the path set by previous administrations. Other commentators anticipated a potential change in the federal government's approach. The May 11 Executive Order proves each of these groups correct, at least in part: the Executive Order builds from existing initiatives by prioritizing the use of the NIST Framework and continuing the emphases of the CNAP, but breaks new ground with potentially industry-altering directives regarding coordination across federal agencies, as well as private sector information-sharing and transparency requirements.

Emphasis on Risk Management

A significant portion of the Executive Order focuses on improving the cybersecurity posture of executive branch agencies. Rather than imposing a prescriptive, one-size-fits-all approach for maintaining and improving this posture, the Executive Order opts for a risk-based approach, requiring the use of the now regularly-used NIST Framework. Under the Executive Order, each executive branch agency must implement risk management measures commensurate "with the risk and magnitude of the harm" that would result from a cybersecurity incident. The intent of the Order is to increase protections against a cyberattack that would result in catastrophic effects on public safety, the economy, or national security.

The NIST framework was specifically intended to be adaptable to any organization, so its use as a tool for federal agencies is a natural progression from existing policy. While there are dozens of other tools available for performing risk assessments, the use of a single tool and uniform structure for assessing risk is intended to facilitate a consistent approach to managing cybersecurity across different federal agencies—a significant change from the status quo.

Each agency must provide a risk management report documenting risk mitigation and acceptance choices to DHS and the Office of Management and Budget ("OMB") within 90 days of the Executive Order. In turn, OMB, in coordination with DHS and with support from the Department of Commerce, will submit a report to the President within 60 days outlining a plan for protecting all executive branch enterprises. The OMB report also will address additional budgetary needs for risk mitigation, and will establish a process for periodic reassessments of cybersecurity gaps.

Modernizing and Consolidating Federal Computer Network Infrastructure

The Executive Order sets the federal government on a path for building and maintaining a more modern IT architecture. The Order directs the American Technology Council—a group of industry representatives established by a different Executive Order dated May 1, 2017[6]—to report on the legal, policy, and budgetary considerations related to transferring all executive branch agencies to "one or more consolidated network architectures" and to utilizing shared IT services. The Executive Order also directs agencies to show preference in their procurement for shared IT services, including email, cloud, and cybersecurity services.

Many observers have cautiously praised this portion of the Executive Order, given the antiquated (and potentially vulnerable) technology on which many federal agencies currently rely. The success of this effort is far from guaranteed, however. Notably, modernizing and potentially unifying the technology used by various agencies would require significant funding, and the Executive Order, which is not a
budgeting document, is silent on that front. In addition, given that the American Technology Council has been in existence for just a few weeks, succeeding in this large-scale, extremely complex, and expensive undertaking will be particularly challenging.

If the Administration were to push agencies toward shared IT services in a significant way, the results could be profound. Each agency now has discretion to determine the extent to which it will utilize such services, and a variety of different approaches are used. Moving to cloud-based solutions would require a re-working of a variety of procurement contract provisions and monitoring mechanisms. It also could result in a massive increase in demand for such services, which could put pressure on available capacity to handle and secure the type of highly sensitive data currently processed and maintained by many federal agencies.

**Cybersecurity of Critical Infrastructure**

The Executive Order includes a number of provisions addressing the cybersecurity of the United States' critical infrastructure. For purposes of the Order, "critical infrastructure" is construed broadly, and includes everything from power generation and electricity transmission to communications facilitation and chemical manufacturing. The Order focuses on both high-level strategy-setting and the development of tactical fixes for several specific threats, although it is primarily an appraisal tool.

As a strategic matter, the Executive Order directs DHS to collaborate with other federal agencies to determine their ability to support the cybersecurity efforts of critical infrastructure entities, a classified group of companies that the U.S. government has identified as being at the greatest risks of cyber attacks that could result in catastrophic effects on public health, economic security, or national security (so-called "section 9 entities"). The agencies also must examine existing federal policies and practices to determine whether they are sufficient to promote "appropriate market transparency" regarding cybersecurity risk management practices, especially by publicly-traded critical infrastructure companies.

On a tactical level, the Executive Order takes aim at automated cyberattacks, prolonged power outages caused by cyber incidents, and cyberattacks on the defense industrial base. The Order directs the Department of Commerce and DHS to jointly lead "an open and transparent process" to identify and spur collaboration around reducing automated cyber-attacks. It directs the Department of Energy to coordinate an assessment of the potential for prolonged power outages associated with cyber incidents. And the Order tasks the Department of Defense with coordinating an assessment of the risks facing the defense industry.

The Order's focus on critical infrastructure underscores a key reality: many infrastructure providers rely on older technology that is particularly vulnerable to attack. And its focus on tactical responses to particular threats builds from actual attacks in the recent past, including the Mirai botnet attack on the domain name service Dyn that took down several prominent websites, as well as a well-publicized attack on several national power grids.

The Order underscores the importance of information sharing and collaboration, including with international partners, to manage cybersecurity risk. It will be important to see whether the Order leads
the government to provide greater incentives for information-sharing by private companies (or penalties for failing to share such information). For years, the federal government has encouraged greater information-sharing by private companies regarding cyberthreats. Eighteen months ago, Congress passed the Cybersecurity Act of 2015, which provided incentives for such sharing. The Department of Justice and the Federal Trade Commission have issued guidance delineating that such sharing, if performed within certain parameters, does not raise antitrust concerns. And the Federal Bureau of Investigation has engaged in concerted information-sharing campaigns through InfraGard and other public-private partnerships. Still, many companies have remained reluctant to share such cyber data based on concerns that the information could be used in future enforcement actions or civil litigation. This calculus may change as the government develops additional policies, although the Executive Order largely continues the same aspirational approach to information sharing.

The Order's focus on the need for market transparency by publicly-traded critical infrastructure providers is also notable, but again simply calls for a report on the sufficiency of existing policies to promote transparency. While awaiting further guidance from this report, companies should carefully consider whether they have undertaken a sufficiently documented risk assessment to protect themselves against second-guessing if an attack is successful, and whether they have adequately disclosed to investors and regulators the nature of the cybersecurity risks they face.

Additional Provisions

Portions of the Executive Order venture beyond addressing the security of federal networks and critical infrastructure. For example, the Order identifies as one of its underlying policies an intent to maintain the U.S. competitive advantage in cybersecurity, and requires certain agencies and departments to provide reports relating to fostering a cybersecurity-skilled workforce in both the public and private sectors. The Order identifies as its other underlying principle: "to promote an open, interoperable, reliable, and secure internet that fosters efficiency, innovation, communication, and economic prosperity, while respecting privacy and guarding against disruption, fraud, and theft." Despite this reference to privacy considerations, the Order's substantive provisions do not expressly provide that agency reports assess or address risks to individual privacy.

Parting Thoughts

The Executive Order's focus on risk-based cybersecurity assessments has implications even for companies that are not directly affected by the Order. When the NIST Framework first was introduced, we predicted that it would become a de facto standard of care. Its adoption as a required tool of all federal executive agencies by the President is a further step toward this result. Companies that have not used the NIST Framework to conduct a well-documented assessment of their cybersecurity vulnerabilities should consider doing so. Relatedly, the Order makes a number of "findings" reflecting its views as to what constitutes "effective risk management," including that it involves regular maintenance and modernization, addressing known vulnerabilities, and consultation with individuals with appropriate expertise. While these are not novel positions, to the extent they reflect administration thinking on appropriate risk management, companies may hear similar views from regulators and enforcement agencies.
The Executive Order primarily requests numerous assessment reports across a broad range of governmental agencies and departments in a relatively short time period. One obstacle to completing the reports, however, is that many cybersecurity leadership positions in key agencies remain unfilled by permanent appointees. What happens next will depend heavily on whether the time periods are sufficient to allow for the development of effective recommendations, and what the reports propose. Potentially affected companies should continue to monitor the news from Washington as the 60- and 90-day deadlines for submitting the reports pass.


The following Gibson Dunn lawyers assisted in the preparation of this client alert: Alexander Southwell, Caroline Krass, Ryan Bergsieker, Jeana Bisnar Maute, and Casper Yen.

Gibson Dunn’s lawyers are available to assist with any questions you may have regarding these issues. For further information, please contact the Gibson Dunn lawyer with whom you usually work or any of the following leaders and members of the firm’s Privacy, Cybersecurity and Consumer Protection practice group.

United States
Alexander H. Southwell - Chair, PCCP Practice, New York (+1 212-351-3981, asouthwell@gibsondunn.com)
Caroline Krass - Chair, National Security Practice, Washington, D.C. (+1 202-887-3784, ckrass@gibsondunn.com)
M. Sean Royall - Dallas (+1 214-698-3256, sroyall@gibsondunn.com)
Debra Wong Yang - Los Angeles (+1 213-229-7472, dwongyang@gibsondunn.com)
Howard S. Hogan - Washington, D.C. (+1 202-887-3640, hhogan@gibsondunn.com)
Joshua A. Jessen - Orange County/Palo Alto (+1 949-451-4114/+1 650-849-5375, jjessen@gibsondunn.com)
Shaalu Mehra - Palo Alto (+1 650-849-5282, smehra@gibsondunn.com)
Karl G. Nelson - Dallas (+1 214-698-3203, knelson@gibsondunn.com)
Eric D. Vandevelde - Los Angeles (+1 213-229-7186, evandevelde@gibsondunn.com)
Benjamin B. Wagner - Palo Alto (+1 650-849-5395, bwagner@gibsondunn.com)
Michael Li-Ming Wong - San Francisco/Palo Alto (+1 415-393-8333/+1 650-849-5393, mwong@gibsondunn.com)
Ryan T. Bergsieker - Denver (+1 303-298-5774, rbergsieker@gibsondunn.com)
Richard H. Cunningham - Denver (+1 303-298-5752, rhcunningham@gibsondunn.com)
Jeana Bisnar Maute - Palo Alto (+1 650-849-5348, jbisnarmaute@gibsondunn.com)

Europe
Ahmed Baladi - Paris (+33 (0)1 56 43 13 00, abaladi@gibsondunn.com)
James A. Cox - London (+44 (0)20 7071 4250, jacox@gibsondunn.com)
Patrick Doris - London (+44 (0)20 7071 4276, pdoris@gibsondunn.com)
Andrés Font Galarza - Brussels (+32 2 554 7230, afontgalarza@gibsondunn.com)
Bernard Grinspan - Paris (+33 (0)1 56 43 13 00, bgrinspan@gibsondunn.com)
Penny Madden - London (+44 (0)20 7071 4226, pmadden@gibsondunn.com)
Jean-Philippe Robé - Paris (+33 (0)1 56 43 13 00, jrobe@gibsondunn.com)
Michael Walther - Munich (+49 89 189 33-180, mwalther@gibsondunn.com)
Nicolas Autet - Paris (+33 (0)1 56 43 13 00, nautet@gibsondunn.com)
Kat Gesing - Munich (+49 89 189 33-180, kgesing@gibsondunn.com)
Update on CFIUS Developments: Proposed Legislation and Reflections on CFIUS Under the Trump Administration

June 23, 2017
CFIUS

We are writing to provide our perspective on recent developments regarding the Committee on Foreign Investment in the United States (“CFIUS”), including the plans of Senator John Cornyn (R-TX) to introduce significant legislation to reform CFIUS, and the operation of CFIUS in the first six months of the administration of President Donald Trump.

Senator Cornyn’s Proposal to Reform CFIUS

Senator Cornyn, the second-highest-ranking Republican Senator, publicly announced on June 22 his intention to introduce a bill to reform the authority and operation of CFIUS. The proposed bill (the “Cornyn Bill”), which Cornyn said would be called the Foreign Investment Risk Review Modernization Act, represents the most significant effort to revise the CFIUS process since the passage of the Foreign Investment and National Security Act of 2007 (“FINSA”), which has governed CFIUS for nearly a decade.

Senator Cornyn’s bill represents the most significant effort to revise the CFIUS process in the last decade.

According to Senator Cornyn, his effort to reform CFIUS is motivated by his view that the Committee is outdated and ill-equipped to handle new threats to U.S. national security, especially from China. In a speech Thursday at the Council on Foreign Relations, the Senator stated that China has sought to “weaponize” investment as part of a strategy to leapfrog the United States’ advantages in technology. He further suggested that there is broad consensus within the U.S. defense and intelligence communities that “gaps” in CFIUS’s authorities present risks to U.S. national security. The proposed bill is intended to fill those gaps.

The concerns driving Senator Cornyn’s reform effort relate in part to a belief within some parts of the U.S. government that China is circumventing CFIUS to gain access to technologies that may have current or future military applications. In particular, Senator Cornyn cited minority investments by Chinese parties in technology start-up companies engaged in developing commercial technologies that may also have defense or intelligence applications. He referred to
a report prepared earlier this year (but not publicly released) by the Defense Innovation Unit Experimental, or DIUx, a component of the Department of Defense, that raised concerns about Chinese investment in early-stage technologies in the areas of robotics, artificial intelligence, automation, and semiconductors, among others.

While we will not know the precise language and effect of the Cornyn Bill until it is introduced, based on Senator Cornyn’s description, the bill would have the following effects:

- **Expand CFIUS’s jurisdiction to review non-control transactions that result in access to technology.** CFIUS currently has jurisdiction to review transactions that result in foreign control of a U.S. business. While the scope of that jurisdiction is extremely broad—extending to minority investments of as little as ten percent equity interest (or less)—it does not currently extend to transactions that result in technology transfers, such as the licensing of intellectual property, that occur outside the context of a controlling investment. The Cornyn Bill apparently would expand CFIUS’s jurisdiction to review such transfers, though it is not clear how broad the authority would be.

- **Expand authority to review overseas joint ventures (“JVs”).** CFIUS already has jurisdiction to review JVs that result in foreign control of U.S. businesses, but Senator Cornyn made clear that he intends for the bill to expand the Committee’s jurisdiction to review JVs that result in technology transfers even where those JVs do not result in control of a U.S. business, such as to JVs outside the United States that involve a transfer of U.S. technology.

- **Increased scrutiny of real estate acquisitions.** Senator Cornyn also suggested that the bill will expand CFIUS’s authority to review real estate transactions. It is unclear exactly what the bill would cover that is not already within CFIUS’s remit, but presumably it would encompass certain real estate outside the context of a transaction resulting in control of a U.S. business.

- **Countries of concern and broader trends.** CFIUS currently reviews transactions on a case-by-case basis, and factors into consideration the threat posed by an acquiror. In that context, CFIUS’s analysis already covers the country of origin for the acquiror, starting with a presumption of approval. Senator Cornyn indicated his bill would add an explicit new country-specific framework that would require the Committee to apply heightened scrutiny to transactions involving certain countries of concern—which he suggested would include China and Russia, but possibly others—that are perceived to present the most serious threat to U.S. national security. He also indicated that he believed the Committee should take into consideration broader trends from the activities of certain countries.

While these would be substantial changes to CFIUS that would materially impact the Committee’s authority and operation, Senator Cornyn expressly rejected other potential reforms:

- **Continued focus on national security; no “net benefit” or “reciprocity” test.** There have been a number of proposals over the years to expand CFIUS’s remit to address economic issues, such as requiring that transactions present a “net benefit” to the United States (akin to the requirements of the Investment Canada Act), or requiring CFIUS to consider whether a foreign investor’s home country would permit a reciprocal investment by a U.S. company in the same industry. Senator Schumer (D-NY), the senior-most Senate Democrat, reportedly is working on such a bill, as is Senator Grassley (R-IA).
Senator Cornyn stated that his bill, by contrast, is “laser focused” on U.S. national security.

- **No expansion to cover greenfield investment.** The Cornyn Bill would maintain CFIUS’s focus on transactions involving existing U.S. businesses. It would not expand the Committee’s jurisdiction to cover startup—or “greenfield”—investments.

- **Voluntary filing framework retained.** Filing with CFIUS currently is voluntary, although the potential consequences of not filing a transaction that later is determined to have an impact on U.S. national security can be quite severe. As such, the process is designed to incentivize parties to notify CFIUS of transactions that could implicate national security interests. The Cornyn Bill apparently will not alter this fundamental framework.

- **No new CFIUS members.** There have been several proposals over the years to add new members to CFIUS. Most recently, Senator Grassley co-sponsored a bill with Senator Stabenow (D-MI) that would give the Secretary of Agriculture and the Secretary of Health and Human Services (who oversees the Food and Drug Administration) permanent seats on CFIUS and would expressly require the Committee to assess the potential impact of transactions on the security of the United States’ food supply. Senator Cornyn stated that his bill would not add new members to CFIUS.

- **No “calling out” specific countries.** While Senator Cornyn made clear that his focus is on China, he also noted that the legislation would not identify any country by name. He suggested the country-specific framework discussed above should be durable and, therefore, it should be up to the Executive Branch, not the Congress, to identify on an ongoing basis countries of concern.

- **No ban on investment.** Senator Cornyn also said that while he is concerned about certain Chinese investments, he believes in preserving an open investment environment, and the bill would not call for a “ban” on any particular investment or from particular countries.

Senator Cornyn indicated that he is putting the finishing touches on the bill and hopes to introduce it soon. He also indicated that it will be a bi-partisan bill that seeks support from Democrats as well as Republicans. We understand that Senator Cornyn has been in discussions with other senators to co-sponsor the bill, and also is working to ensure the introduction of a companion bill in the U.S. House of Representatives.

To be sure, there are strong potential challenges to passing any significant legislation. In the most bi-partisan of times, it is not easy to move legislation; the current environment in Washington is, to say the least, something less than bi-partisan. Plus, other issues wholly unrelated to investment or CFIUS are occupying, and will continue to occupy, Congressional attention for the rest of this year.

But there are also reasons to believe that the bill (or a version of it) may ultimately become law. Senator Cornyn is among the most influential members of the Senate, and there appears to be an emerging bi-partisan consensus that CFIUS should be strengthened, principally to address growing concerns about Chinese investment. Senator Cornyn also has been discussing the bill with key members of the Trump Administration, including Treasury Secretary Steven Mnuchin. Within the last several weeks, all the leaders of the U.S. intelligence community, as well as Secretary of Defense Mattis, Secretary of Commerce Ross, and Secretary of the Treasury Mnuchin, have expressed support for CFIUS reform. These comments from Administration
officials indicate potential momentum behind Senator Cornyn’s efforts to reform CFIUS. For these reasons, we believe Senator Cornyn’s effort merits close attention.

Within the last several weeks, all the leaders of the U.S. intelligence community, as well as Secretary of Defense Mattis, Secretary of Commerce Ross, and Secretary of the Treasury Mnuchin, have expressed support for CFIUS reform. These comments from Administration officials indicate potential momentum behind Senator Cornyn’s efforts to reform CFIUS. For these reasons, we believe Senator Cornyn’s effort merits close attention.

Reflections on CFIUS Under the Trump Administration

Senator Cornyn’s announcement comes at a time of dramatic change in the composition of foreign investment in the United States and also in the early days of a new presidential administration that in most respects marks a radical change from the administration of former President Barack Obama. These shifts are impacting CFIUS itself as well as parties who bring transactions to the Committee for review.

With respect to the composition of the investments that CFIUS reviews:

- CFIUS’s caseload has increased considerably, even over what likely was a record year in 2016. Through today, CFIUS has accepted notices of over 120 transactions in 2017, on pace to set an all-time record; by comparison, the Committee received notices of just 97 transactions in all of 2013. This increase is the result of a variety of factors, including strong merger and acquisition activity in general, greater awareness of CFIUS among transaction parties, more aggressive efforts by CFIUS in pursuing certain non-notified transactions, and more transactions involving investors from countries that tend to receive greater scrutiny from CFIUS.

- There now is a much greater proportion of investment from developing markets, especially China. While the numbers of CFIUS cases and specific countries from 2015 and 2016 are not available, our estimation is that CFIUS likely reviewed as many as 80 to 90 Chinese transactions in 2016, as compared with just four Chinese transactions in total from 2005 to 2007.

- Compared with several years ago, the Chinese investment activity is more diverse, covering more sectors, with varied transaction structures, and is often multi-national in scope.

These factors result in an extremely busy and challenging docket for CFIUS at a time when the Trump Administration still is filling many of the key positions responsible for overseeing CFIUS.
The effect is that CFIUS currently is lacking clear policy direction and is more risk adverse, resulting in a harder road to approval for some transactions.

We expect the next six months, at a minimum, to be a very challenging time for CFIUS transactions, particularly those from China.

In sum, since President Trump’s inauguration, we have seen a reduced willingness on the part of CFIUS to attempt to resolve national security risks through negotiated mitigation, and a greater ability among the security agencies to obtain consensus decisions within the Committee to simply prohibit challenging transactions. This tougher line likely is being driven by multiple factors, including resource constraints arising from the sheer volume of transactions under consideration, greater risk aversion among agencies driven in part by continuing widespread vacancies in senior political positions (and a resulting lack of political direction and accountability), and more significant national security issues presented by some transactions. We expect the next six months, at a minimum, to be a very challenging time for CFIUS transactions, particularly those from China.

At the same time, we do not yet see a fundamental change in the policy underlying CFIUS’s consideration of cases. Thus, notwithstanding vigorous anti-trade rhetoric during the campaign and positions held by some senior Administration officials, there is not yet an affirmative departure from the United States’ longstanding policy of openness to foreign investment. In addition, we are cautiously optimistic that as additional senior policy positions are filled, firmer views about the overall approach to investment will take hold, and the CFIUS officials will be better equipped to engage in a more timely manner in transactions, and to seek solutions to national security concerns. The Congressional debate, however, around China and the prospect for CFIUS legislation bears monitoring, as it could significantly shape the United States’ approach to investment from China in certain areas.

The Congressional debate, however, around China and the prospect for CFIUS legislation bears monitoring, as it could significantly shape the United States’ approach to investment from China in certain areas.

* * *
We hope that you find this analysis useful. Please do not hesitate to contact the following members of our CFIUS practice if you would like to discuss any aspect of the foregoing in further detail:

Mark Plotkin  
+1 202 662 5656  mplotkin@cov.com

David Fagan  
+1 202 662 5291  dfagan@cov.com

Stuart Eizenstat  
+1 202 662 5519  seizenstat@cov.com

Alan Larson  
+1 202 662 5756  alarson@cov.com

Peter Lichtenbaum  
+1 202 662 5557  plichtenbaum@cov.com

John Veroneau  
+1 202 662 5034  jveroneau@cov.com

Roger Zakheim  
+1 202 662 5959  rzakheim@cov.com

Damara Chambers  
+1 202 662 5279  dchambers@cov.com

Heather Finstuen  
+1 202 662 5823  hfinstuen@cov.com

Meena Sharma  
+1 202 662 5724  msharma@cov.com

Jonathan Wakely  
+1 202 662 5387  jwakely@cov.com

This information is not intended as legal advice. Readers should seek specific legal advice before acting with regard to the subjects mentioned herein.

Covington & Burling LLP, an international law firm, provides corporate, litigation and regulatory expertise to enable clients to achieve their goals. This communication is intended to bring relevant developments to our clients and other interested colleagues. Please send an email to unsubscribe@cov.com if you do not wish to receive future emails or electronic alerts.
To Our Clients and Friends:

In honor of Data Privacy Day--an international effort to raise awareness and promote privacy and data protection best practices--we offer this fifth edition of Gibson Dunn's Cybersecurity and Data Privacy Outlook and Review. In 2016, companies, governments, and consumers were again challenged to navigate an evolving landscape of cybersecurity and privacy issues. This year saw flash points impacting the trajectory for data breach litigation, the future for privacy class actions, and the scope of government powers to both regulate data collection practices and gather data itself. Cybersecurity also burst onto the international regulatory and political scene.

Among other developments, this year the Supreme Court issued its decision in *Spokeo, Inc. v. Robins*, a long-awaited development addressing (somewhat) plaintiffs' burden to show concrete injury to satisfy Article III standing. Plaintiffs and defendants had argued for years over what allegations are sufficient to show a true privacy harm. While it may not have resolved all the open issues, the *Spokeo* decision has already been cited over one thousand times.

In addition, plaintiffs pursued new avenues for litigation opened by new technologies, including the use of biometric information and connected devices, and new theories for established standbys. The year saw a number of resolutions of several closely watched data breach class actions and cybersecurity-related shareholder derivative suits. A number of additional regulatory agencies entered the privacy game, either bringing enforcement actions or issuing privacy/cybersecurity guidance. This year the government also found itself on the other side of the privacy debate, in legal battles over its ability to collect personal information from companies (i) without notice to the subject, and/or (ii) outside the United States. And, of course, EU and U.S. regulators agreed to a new framework for international data transfers--the Privacy Shield--which has already seen its first legal challenges.

We cover these topics and many more in this year's Review: (i) civil litigation; (ii) U.S. government regulation of privacy and data security; (iii) U.S. government data collection; and (iv) international developments. For additional coverage of international developments, please see our separate International Cybersecurity and Data Privacy Outlook and Review.
# Table of Contents

## I. Civil Litigation

A. Standing After *Spokeo*

B. Data Breach Litigation
   1. Litigation
   2. Data Breach Settlements
   3. Shareholder Derivative Suits

C. Interceptions and Eavesdropping
   1. Email Scanning
   2. Call Recording
   3. Other "Interceptions"

D. Telephone Consumer Protection Act

E. Video Privacy Protection Act

F. California's Song-Beverly Credit Card Act and Point-of-Service Data Collection

G. Biometric Information Privacy Acts

H. Internet of Things and Device Hacking

I. Cybersecurity Insurance

## II. U.S. Government Regulation of Privacy and Data Security

A. Enforcement and Guidance
   1. Federal Trade Commission ("FTC")
   2. Department of Health and Human Services ("HHS")
   3. Securities and Exchange Commission
   4. Federal Communications Commission ("FCC")
In Spokeo, Inc. v. Robins, the Supreme Court considered whether a statutory violation, without resulting "concrete" injury, satisfies the "injury-in-fact" requirement of Article III.[1] On May 16, 2016, the Court issued its much-anticipated ruling: that a plaintiff must suffer an injury in fact that is both particularized and concrete to have standing to sue, and that "a bare procedural violation, divorced from any concrete harm" to the plaintiff, cannot satisfy this injury-in-fact requirement.[2] The decision disapproved of many lower court decisions that deemed an alleged statutory violation as sufficient, on its own, to satisfy the standing requirements of Article III--and thus Spokeo was poised to have broad ramifications across the privacy landscape. In practice, however, lower courts' interpretation and application of Spokeo has been mixed. While defendants can certainly claim a nominal victory--merely alleging a statutory violation will not necessarily confer standing post-Spokeo--the decision did not provide the clarity that either plaintiffs or defendants had desired.
1. The Spokeo Decision

Thomas Robins filed a class action lawsuit against Spokeo, the operator of a "people search engine," alleging violations of the Fair Credit Reporting Act of 1970 ("FCRA"). Robins claimed that Spokeo "willfully failed to comply" with the FCRA because Spokeo generated a profile that contained inaccurate information about him. The district court dismissed the suit for lack of standing on the ground that Robins had not "properly pled" an injury in fact. The Ninth Circuit subsequently reversed, finding that Robins' allegations were sufficient to satisfy the injury-in-fact requirement since "Spokeo [had] violated his statutory rights, not just the statutory rights of other people," and his "personal interests in the handling of his credit information are individualized rather than collective."

On appeal, the Supreme Court found that the Ninth Circuit's analysis was "incomplete," because the injury-in-fact element requires a plaintiff to allege that an injury is both particularized and concrete, and the Ninth Circuit had "overlooked" the concrete factor. The Supreme Court accordingly vacated the Ninth Circuit's opinion and remanded the case to the Ninth Circuit to complete the concrete injury inquiry.

Although the Supreme Court did not take a position as to whether Robins had established a concrete injury, it provided guidance on how to conduct the analysis: A concrete injury need not necessarily be "tangible," but it must be "real" and not merely "abstract" in nature. In determining whether an intangible injury is concrete, the Court noted that it is "instructive" to consider both the "judgment" of Congress, as well as whether the alleged injury "has a close relationship to a harm that has traditionally been regarded as providing a basis for a lawsuit." At the same time, the Court cautioned that "Congress' role in identifying and elevating intangible harms does not mean that a plaintiff automatically satisfies the injury-in-fact requirement whenever a statute grants a person a statutory right and purports to sue to vindicate that right." Put differently, "Article III standing requires a concrete injury even in the context of a statutory violation." The Court also noted that a plaintiff cannot "allege a bare procedural violation, divorced from any concrete harm, and satisfy the injury-in-fact requirement of Article III." The Court found it "difficult to imagine" how dissemination of certain technically inaccurate information (such as an improper zip code), without more, could constitute a sufficiently concrete harm to satisfy standing.

2. Post-Spokeo Standing Decisions in Privacy Cases

In the months since the Spokeo decision, courts have issued more than one hundred decisions addressing Spokeo's impact on the question of plaintiffs' Article III standing in privacy-related cases. The majority of those decisions have been rulings on defendants' motions to dismiss claims based on the FCRA, Fair Debt Collections Practices Act ("FDCPA"), Fair and Accurate Credit Transactions Act ("FACTA"), and Telephone Consumer Protection Act ("TCPA"). The majority of those decisions found standing to be sufficiently alleged, but a significant percentage found standing to be lacking. Moreover, courts addressing other privacy-related claims—including the Cable Communications Privacy Act ("CCPA"), the Truth in Lending Act, the Video Privacy Protection Act ("VPPA"), and the Electronic Funds Transfer Act—collectively found standing to be insufficiently alleged approximately as often as they found it to be sufficient. Indeed, there are decisions going different ways for almost every legal claim and in each
of the key privacy-related factual contexts. Here we briefly address post-*Spokeo* standing decisions in a number of key privacy-related areas.

**Data Breach.** Both the Sixth and Seventh Circuits issued decisions this year finding that data breach subjects had standing to sue based on alleged fraud-prevention expenses and substantial risk of harm from identity theft.[15] On the other hand, a number of district courts dismissed claims for lack of standing where the threat of future harm from a data breach was too speculative.[16] Yet other courts assessed the standing arguments on a plaintiff-by-plaintiff and injury-by-injury basis, permitting those alleging actual instances of identity theft and dismissing others.[17] In July, the Third Circuit heard oral argument for an appeal wherein plaintiffs argued that the mere statutory violation of the FCRA's requirement to protect data is sufficient to confer standing and that nefarious data breaches always carry imminent risk of harm and automatically confer standing. Mere days before publication, the court ruled in plaintiffs favor, vacating the lower court's dismissal.[18]

**Unlawful Disclosure.** The Third Circuit issued a decision this year finding that plaintiffs alleging unlawful disclosure of legally protected information, specifically web tracking data, in violation of several statutes, including the VPPA, the Wiretap Act, and the Stored Communications Act ("SCA"), alleged a concrete injury sufficient to confer standing.[19] Some district courts used the same reasoning to justify analogous rulings based on allegedly unlawful disclosure under FDCPA,[20] while other district courts reached the same conclusion on more tangible harms (e.g., risk of identity theft) for disclosures allegedly in violation of the Song-Beverly Consumer Warranty Act, Cal. Civ. Code § 1790 *et seq.*[21] However, yet other courts reached the opposite conclusion for disclosures allegedly in violation of FACTA, finding that plaintiffs had alleged insufficient risk of the harm envisioned by Congress.[22]

**Unlawful Retention.** Unlawful retention cases have trended in defendants' favor. Courts have found that unlawful retention of data alone--whether an alleged violation of the CCPA, the Biometric Information Privacy Act ("BIPA"), or state analogs to the VPPA--is insufficient to establish standing.[23] In ruling on a motion to dismiss claims under the CCPA, the Eighth Circuit explained that mere retention is a "bare procedural violation."[24] To sufficiently plead concrete and particularized injury, plaintiffs would have had to identify some improper use of, or harm flowing from, the retention of the information in question.

**Unlawful Acquisition/Use.** In *Matera v. Google Inc.*, an email scanning case discussed further herein, the Northern District of California ruled that alleged interceptions of communications without consent, in violation of the Wiretap Act and state law analogs, constitute injury in fact sufficient to satisfy standing.[25] However, courts have also found that unlawful acquisition of other types of information—for example, requesting customers' zip codes in violation of state law--does not automatically confer standing.[26]

**TCPA Claims.** Rulings in TCPA cases have also gone both ways, but have trended in plaintiffs' favor, finding that telemarketing activities and unsolicited automated calls prohibited by the statute infringe the precise interests Congress sought to protect, and finding that such allegations are concrete violations of substantive rights and therefore confer standing.[27] Other courts relied on more conventional grounds
to reach the same conclusion, allowing plaintiffs to move forward with more tangible, but very minimal, injury.[28] However, some courts have dismissed TCPA cases where the alleged injury was insufficient or not tied to a particular violation.[29]

3. Looking Ahead

The Supreme Court's decision to remand the Spokeo case to the Ninth Circuit without addressing the alleged injury in that case provided little guidance concerning what constitutes a concrete injury, leaving ample room for arguments to be made on both sides, as illustrated above.[30] Many feel that, in doing so, the Court created an unpredictable new legal landscape.[31] Spokeo did not provide a bright-line rule squarely prohibiting plaintiffs from suing for intangible injuries (such as the publication of inaccurate information on the internet).[32] However, it has been interpreted to prevent plaintiffs from suing for pure procedural violations and from relying purely on alleged statutory violations, in the absence of actual harm. Spokeo's most profound impact thus may be in class actions where plaintiffs seek statutory damages on a class-wide basis without proof that any class members actually suffered actual injury.[33]

In the oral argument on remand before the Ninth Circuit on December 13, 2016, the parties in Spokeo presented precisely these issues. Plaintiff argued that the court should assess the type of harm alleged, and whether it goes to the concrete interest that Congress intended to protect. Defendant argued that the court must assess the specific harm alleged for the specific plaintiff, and may only find standing to be satisfied if the particular plaintiff has plausibly alleged specific particularized and concrete harm that is traceable to the actual challenged conduct of the defendant. While the decision will not be binding precedent in other circuits, many will be watching to see how the Ninth Circuit interprets the law.

B. Data Breach Litigation

Like 2015, 2016 had its share of massive data breaches. And, again, a global study revealed that both the number and the cost of breaches increased.[34] A report from the California Attorney General reported that from 2012 to 2015, the number of data breaches reported to the California Attorney General also rose from 131 to 178.[35] The trajectory continued in 2016, with a 40% increase in data breaches reported to the New York Attorney General by May.[36] The average data breach cost for U.S. organizations was $7.01 million, a 20% increase since 2014.[37]

As in previous years, litigation has followed closely on the heels of almost every large breach. Below, we review a number of high-profile breach announcements this year, the litigation that followed, and key developments in ongoing data breach litigation, including important decisions and settlement trends.

1. Litigation

a. High-Profile Breaches in 2016

Major data breaches in 2016 impacted a number of different industries and involved a number of different types of personal information. Hackers targeted customer login information, payment information, and employees' personal information, among others. In many instances, though not all, litigation quickly followed the announcement of the breach.
i. Election-Related Hacks

In 2016, the U.S. government publicly blamed Russia for a July cyberattack on the Democratic National Committee.[38] The DNC hack resulted in thousands of internal emails being publicly posted online, including through the website WikiLeaks. In October, emails from Hillary Clinton's campaign chairman John Podesta began appearing online, and a private security firm attributed the hack to Russian foreign intelligence activity occurring months earlier.[39] Subsequently, the CIA concluded that government-sponsored Russian hackers were responsible for the Podesta hack.[40] U.S. intelligence agencies also concluded that the Russian government additionally attempted to hack the Republican National Committee.[41]

In late December, President Obama announced sanctions against Russia's intelligence apparatus, including expelling 35 individuals for spying while posing as diplomatic officials, shutting down two Russian compounds, and sanctions on Russian intelligence agencies, top Russian intelligence officials, and three companies and organizations allegedly involved in the hacking.[42]

ii. Login Information

LinkedIn. Professional networking website LinkedIn saw fallout from a previous loss of user login information. LinkedIn disclosed in May 2016 that 100 million usernames and passwords from a 2012 breach had been posted online.[43] In 2015, LinkedIn had settled class action litigation over the breach for $1.25 million.[44]

iii. Health Information

21st Century Oncology. Following an October 2015 data breach that allegedly disclosed the sensitive identifying and medical information of 2.2 million patients, 21st Century Oncology, a network of national cancer treatment centers, was hit with numerous class actions starting in March 2016.[45] The MDL Panel consolidated and transferred sixteen cases against 21st Century Oncology to the Middle District of Florida for combined proceedings.[46] Preliminary litigation involving case management and appointment of counsel is ongoing.

iv. Payment Information

Kimpton. In September, boutique hotel chain Kimpton announced that hackers may have obtained the information from credit and debit cards used in more than 60 of its hotels and restaurants between February and July 2016.[47] After Kimpton's data breach was announced on September 1, a federal class-action lawsuit followed within a month, alleging breach of implied contract, violation of California unfair business practices laws, and negligence.[48] Kimpton filed a motion to dismiss in December, alleging that the plaintiff's claims of "increased risk" of identity fraud, "loss of privacy," and "deprivation of the value of personal information" did not give rise to standing, and also arguing that the plaintiff failed to state a claim for his contractual or state competition law claim.[49]

Wendy's. In January 2016, fast-food restaurant chain Wendy's notified customers that a malware attacker obtained payment card information from 300 of its stores. In July, Wendy's revised its
announcement to say that there were two malware attacks and they impacted 1,025 stores.[50] Litigation on behalf of a consumer class followed. In July the court dismissed the consumer class action without prejudice for lack of Article III standing.[51] The named plaintiff used his debit card at Wendy's during January 2016 and subsequently experienced two fraudulent charges, for which his bank reimbursed him.[52] The court found that the plaintiff did not suffer any monetary harm from the unauthorized charges, and therefore could not allege "actual harm sufficient to establish injury-in-fact."[53] Plaintiff also alleged that he and members of the proposed class were at continuing risk of harm for identity theft and identity fraud, but the court found that the threat of future harm from identity theft is "highly speculative," and therefore cannot form the basis for standing.[54] Plaintiff filed an amended complaint shortly afterwards, alleging breach of implied contract, negligence, and violations of state consumer protection and data breach notification statutes.[55] Wendy's again filed a motion to dismiss, which the court has not yet ruled on.[56]

A class of financial institutions also seeks damages for costs associated with cancelling and reissuing cards, notifying consumers of the breaches, refunding fraudulent charges, and increasing their monitoring activity.[57] Plaintiffs alleged causes of action for negligence and negligence per se arising from defendant's failure to maintain adequate cybersecurity measures in violation of the Federal Trade Commission Act.[58] Wendy's filed a motion to dismiss, alleging plaintiffs failed to state their claims.[59] The motion is fully briefed and pending a ruling.

In December, a Wendy's shareholder brought a derivative suit for breach of fiduciary duty, alleging that the board of directors and executive officers failed to protect Wendy's payment system and did not disclose the data breach until after a report from a third-party security researcher.[60] Wendy's has not yet responded to the complaint.

v. Law Firms and Business Information

In March, hackers breached multiple large U.S. law firms, including Cravath Swaine & Moore LLP and Weil Gotshal & Manges.[61] The U.S. Attorney's Office for the Southern District of New York and the FBI then investigated whether any of the stolen information had been used for insider training. Although Weil Gotshal declined to comment, Cravath confirmed it had suffered a "limited breach" of its network in 2015 but stated it was not aware of any improper use of the information.[62] In December, three Chinese nationals were indicted for the hack, and prosecutors said the hackers had used the stolen information to commit $4 million in securities fraud.[63] In addition to the federal charges, the SEC filed a parallel civil enforcement action, which included a request to freeze the hackers' assets.[64]

vi. Employee Information

In several breaches, employees alleged that their employers, or third parties working at their employers' behest, compromised the employees' personal information.

Sprouts Farmers Market and Seagate Technology. Sprouts Farmers Market and Seagate Technology faced similar lawsuits alleging that the companies compromised employees' Form W-2 data by sending it to cybercriminals through a "phishing" scam.[65] The Sprouts case is in the preliminary case management stages following MDL transfer and a consolidated complaint has not yet been
filed.[66] The consolidated class action against Seagate, which alleges negligence, breach of implied contract, and breach of California consumer protection law,[67] is currently stayed through the end of January 2017 at the parties’ request.[68]

**Lamps Plus.** Lamps Plus is currently defending a putative class action filed by an employee (Frank Varela) who claimed employees' data had been exposed through a payroll provider.[69] Plaintiff alleged statutory claims under California's Consumer Records Act, California's Unfair Competition Law, the FCRA, and various common law claims.[70] Lamps Plus moved to compel individual arbitration based on an arbitration agreement the employee signed as a condition of employment.[71] Instead, the district court authorized class-wide arbitration (and accordingly dismissed the class action complaint).[72] Lamps Plus appealed to the Ninth Circuit,[73] and additionally moved to stay the class arbitration pending the outcome of its appeal.[74] The district court denied the motion to stay on December 27, 2016.[75] Briefing in the appeal is due in February 2017.[76]

**b. Update on Major Data Breach Cases from Prior Years**

Many data breach cases from 2015 headed for settlement instead of progressing to resolution on the merits or even to class certification, as discussed in detail in the Settlements section below. However, some cases did have significant rulings before settlement, and others continue to wind their way through the courts.

**i. New Litigation in Previous Breaches**

**Sony Pictures Entertainment.** After a 2014 cyberattack, Sony employees sued the company for the disclosure of their personal information, and the parties reached a final settlement in April 2016, discussed further below in Section II.B.2.c.[77] In July 2016, Sony found itself the target of another lawsuit related to the same data breach, this time from a movie producer.[78] Possibility Pictures II brought a claim for breach of contract, arguing that Sony breached its distribution agreement by failing to put in place basic cybersecurity measures that would have prevented the breach, and that Sony was responsible for the loss in revenue when the producer's film was pirated following the breach.[79] Sony filed a motion to compel arbitration and stay the proceedings, stating that a clause in the distribution agreement requires the breach of contract claims to be settled in arbitration.[80] The court has not yet ruled on whether to require arbitration.

**ii. Ongoing District Court Litigation**

**Anthem.** In February 2015, Anthem, the nation's second-largest health insurer, announced that hackers had accessed a database containing approximately 80 million customer records, including names, birthdates, and Social Security numbers.[81] More than 100 breach-related class actions filed against Anthem were consolidated in a single multidistrict litigation in the Northern District of California.[82] In February 2016, the court granted in part and denied in part motions to dismiss plaintiffs' amended complaint.[83] The court ruled that (1) the loss of "benefit of the bargain" concerning personal information—that plaintiffs' would not have entrusted their personal information to the insurer if they knew the insurer had "inadequate" safeguards in place—constitutes harm under New York's General Business Law (GBL), and (2) the loss of value of personally identifiable information is
a cognizable economic injury under the GBL.[84] The court also found that California's Uniform Commercial Law allows plaintiffs to seek restitution for profits that Anthem allegedly gained by providing "lax security measures."[85] Fact discovery in the case closed on December 1, 2016. The deadline to file for class certification is March 10, 2017.[86]

Ashley Madison. Following the 2015 public disclosure of account information from Ashley Madison, a website advertised as a place for married individuals to arrange extramarital liaisons, plaintiffs in the ensuing data breach litigation filed a motion for leave to proceed under pseudonyms,[87] which the court denied in part, finding that the users seeking to serve as class representatives in the multidistrict litigation must be publicly identified.[88] In August 2016, Ashley Madison's parent organization, Avid, brought a motion to dismiss or to compel arbitration pursuant to the Ashley Madison website terms and conditions.[89] The motion is pending.

iii. Appellate Litigation

Horizon Healthcare. The Third Circuit recently considered an appeal of a data breach class action arising from the theft of two laptops containing identifying, demographic, and medical information on almost 840,000 Horizon Healthcare policyholders in 2014.[90] Horizon Healthcare policyholders brought a class action under the FCRA and state law claiming "economic damages and other actual harm," which was dismissed in 2015 for a lack of standing.[91] The district court found plaintiffs had not shown actual harm as a result of the breach, because they had not suffered actual economic injury, and an increased likelihood of future harm was insufficient to create standing.[92] The plaintiff class appealed to the Third Circuit, arguing that Horizon violated the FCRA's requirement to protect personal information, and that this statutory violation qualifies as injury-in-fact for Article III standing.[93] Plaintiffs also argued that a "malicious and sophisticated" data breach confers Article III standing on victims, because the nefariousness of the breach means that harm to plaintiffs is imminent.[94] The Third Circuit reversed on January 20, 2017, vacating the dismissal and finding that "[e]ven without evidence that the plaintiffs' information was in fact used improperly, the alleged disclosure of their personal information created a de facto injury."[95]

The Home Depot. Between April 2014 and September 2014, hackers stole the personal and financial information of up to 56 million Home Depot customers.[96] Following this breach, consumers and financial institutions sued Home Depot, and the actions were consolidated through the MDL Panel into a consumer class action and a financial institution class action.[97] The financial institution class members issued and owned compromised payment cards, and alleged negligence, negligence per se, injunctive and declaratory relief, and state statutory violations under Alaska, California, Connecticut, Florida, Illinois, Massachusetts, Minnesota, and Washington laws.[98]

In May 2016, the court ruled on Home Depot's motion to dismiss the financial institution class action and allowed nearly all the financial institutions' claims to proceed.[99] In July, Home Depot sought an order for immediate interlocutory appeal to the Eleventh Circuit of the order refusing to toss the data breach claims.[100] The questions raised on appeal include whether banks have Article III standing to assert claims arising out of a data breach, and whether retailers owe banks a duty to protect against third-
party hacks.[101] The district court has not yet ruled on whether to certify the appeal. The parties are currently conducting discovery, and plaintiffs' motion for certification is due January 30, 2018.[102]

On August 23, 2016, the court granted final approval of the consumer class action settlement, which is discussed further below in Section II.B.2.a.

c. Decisions in Data Breach Cases in 2016

i. Standing

Standing was a major issue in data breach litigation throughout 2016, especially after May, as courts grappled with the Supreme Court's decision in *Spokeo, Inc. v. Robins*.[103] The Seventh Circuit continued its recent trajectory, finding again that data breach victims had standing to sue, and the Sixth Circuit issued a decision following the Seventh Circuit's example. However, district courts in a number of other circuits dismissed data breach claims where the alleged harm was simply too speculative.

The Seventh Circuit revived a data breach suit after a lower court dismissed for failure to allege actual injury in *Lewert v. P.F. Chang's China Bistro, Inc.*[104] An Illinois district court ruled in 2014 that a proposed class of P.F. Chang's customers suing the restaurant chain over a data breach could not pursue their breach of implied contract and Illinois consumer protection claims.[105] The lower court held that plaintiffs failed to allege successful fraudulent charges, and stated that "speculation of future harm does not constitute actual injury."[106] The Seventh Circuit reversed, holding that the plaintiffs had standing because they alleged spending time and money monitoring their financial information to protect against unauthorized charges and identity theft.[107] In doing so, the Seventh Circuit relied on its July 2015 ruling in *Remijas v. Neiman Marcus Group, LLC*[108] that customers whose credit card information has been stolen in a data breach have standing to sue not only after they are hit with fraudulent charges, but also for fraud-prevention expenses such as credit monitoring.[109]

The Sixth Circuit followed suit in *Galaria v. Nationwide Mutual Insurance Company.*[110] The lawsuit, claiming invasion of privacy, negligence, bailment, and violations of the FCRA, was brought by consumers after Nationwide experienced a data breach allegedly involving plaintiffs' personal information.[111] The district court had dismissed plaintiffs' claims, holding that alleged harms including a heightened risk for fraud and paying for mitigation costs such as credit freezes were not sufficient to establish standing.[112] The Sixth Circuit reversed,[113] holding that plaintiffs' alleged injury was sufficient under *Spokeo*, stating that "allegations of a substantial risk of harm, coupled with reasonably incurred mitigation costs, are sufficient to establish a cognizable Article III injury at the pleading stage of the litigation."[114] The Sixth Circuit further noted that its holding was consistent with the recent Seventh Circuit decisions in *Remijas* and *Lewert*.[115]

However, at the district court level, judges regularly dismissed data breach complaints because the threat of future harm to consumers was not sufficient to sustain Article III standing. In *Khan v. Children's National Health Systems*, the putative class action plaintiff alleged that Children's had violated Maryland and D.C. consumer protection laws by failing to adequately protect personal data compromised in a breach.[116] The court found that these allegations amounted to "bare procedural harm" under *Spokeo*, and that because plaintiff "failed to connect" them to a "concrete harm," she failed to establish
In Cox v. Valley Hope Association, the court found that a heightened risk for future identity theft after an unencrypted laptop with personal patient information was stolen was too speculative to constitute concrete harm. Other courts followed similar logic, finding that a continuing risk of harm for identity theft and identity fraud is "highly speculative" and not "certainly impending." One court ruled that without evidence that stolen information had been used to commit any identify theft, fraud, or another act that resulted in harm to any plaintiff, plaintiffs did not have standing.

Another district court took a more moderate approach in the In re Zappos.com, Inc. lawsuit, which stems from a 2012 data breach of an online retailer's servers containing the personally identifying information of approximately 24 million customers. Rather than dismiss the complaint outright, the district court narrowed the multidistrict litigation--dismissing thirteen plaintiffs for lack of Article III standing and dismissing several causes of action--but allowed other plaintiffs leave to amend several claims. The dismissed plaintiffs claimed that their email accounts were "accessed by hackers and used to send unwanted advertisements to people in [their] address book[s]" and that the hackers' unauthorized access had devalued their personal information, but the court found these claims too conjectural to allege actual injury as required by Article III. However, where plaintiffs alleged instances of actual identity theft and fraud, the court found those allegations sufficient to establish standing.

### ii. Negligence and the Economic Loss Doctrine

At the circuit level, the Third Circuit rejected a Pennsylvania data breach class action because Pennsylvania's economic loss doctrine--which provides that "no cause of action exists for negligence that results solely in economic damages unaccompanied by physical injury or property damage"--barred the suit. Plaintiffs alleged that they had suffered damages in early 2015, when unknown third parties breached Benecard's computer system and gained access to plaintiffs' personal and confidential information. Plaintiffs claimed they suffered financial harm when these unknown third parties used plaintiffs' information to file fraudulent tax returns and the IRS issued tax refunds to the unknown third parties. Because the plaintiffs were "not in contractual privity with Benecard and thus ha[d] no contractual remedy," they brought claims based on theories of negligence. The Third Circuit ruled that Pennsylvania's economic loss doctrine, which "generally precludes recovery in negligence actions for injuries which are solely economic," barred the suit.

### iii. "Highly Offensive" Invasions of Privacy

A class of Barnes & Noble customers whose credit and debit card information had been compromised in a 2012 incident involving PIN pad terminals were found to have standing, but the court still dismissed the case, finding plaintiffs failed to adequately plead their claims. The court held, among other things, that the personal information disclosed in this case--specifically, payment card information, personal identification numbers, and names--did not qualify as "private facts, the disclosure of which would be highly offensive to a reasonable person" under Illinois law, as required for an invasion of privacy claim.
2. Data Breach Settlements

As more and more data breach cases are surviving motions to dismiss on standing grounds, one might have expected to see some helpful guidance from the courts on the key issues in data breach cases -- cybersecurity safeguards and protocols, breach readiness, quality of remediation, timing and quality of breach notice, causation and quantification of harm, etc. Instead, most major data breach cases have settled on a class-wide basis. This is unsurprising in light of the dynamics that are present in almost every data breach case: data breaches are singular events for the victim company; they attract regulatory attention and bad press; and they have negative reputational and branding impacts for as long as they are remembered.

In 2016, a number of defendants chose this route and settled major breach cases on a class-wide basis. In particular, three high-profile breaches--Home Depot, Target, and Sony--all were resolved (at least in part) through class-wide settlement. In the Target litigation, both the consumer class and the class of financial institutions ended in settlements. As discussed in further detail below, the components of these settlements are familiar when viewed in the broader context of historical settlements. They consist of funds for class claims for different kinds of losses, credit monitoring services, and reform of security-related practices (including training, disclosures, program design and oversight, as well as vendor management). However, the funds and attorneys' fees are considerably larger than in previous years. Of perhaps most interest, the Target settlement of financial institution claims provided a window into the allocation of post-breach costs as between the breach-victim merchant on the one hand, and the financial institutions and card networks incurring costs on the other hand.

a. Home Depot

After a 2014 data breach, plaintiffs brought a class action lawsuit on behalf of approximately 56 million Home Depot customers whose payment or contact information was implicated. The parties fully briefed a motion to dismiss the claims but moved to preliminarily approve a proposed settlement before the court issued an order.[133] On August 23, 2016, Judge Thomas Thrash of the Northern District of Georgia granted final approval of the class action settlement.[134] The settlement requires Home Depot to pay up to $13 million into a fund to compensate class members (up to $10,000 each) for documented out-of-pocket losses and consequential expenses (including time spent remedying issues related to identity theft), $6.5 million to fund 18 months of identity protection services, and to implement certain data security measures in U.S. stores for two years.[135] The required security measures include creation of a security officer position, security assessments and enhanced safeguards, security-related disclosures for customers, employee training, and ensuring that The Home Depot's vendors maintain similar practices.[136] The court ordered Home Depot to pay $1,000 for each representative plaintiff and $7.536 million in attorneys' fees.[137] The claims brought against Home Depot by a putative class of financial institutions survived a motion to dismiss,[138] and the parties are currently conducting discovery. The financial institution plaintiffs' motion for certification is due January 30, 2018.[139]
b. Target

Following a 2013 data breach, plaintiffs brought a lawsuit on behalf of approximately 110 million customers of Target Corporation ("Target") whose payment or contact information was implicated. The parties conducted considerable discovery and settlement negotiations while waiting for a decision on Target's motion to dismiss.[140] The parties then signed a class-wide settlement agreement less than one month following the court's decision denying the motion to dismiss in part,[141] and on November 17, 2015, Judge Paul Magnuson of the District of Minnesota granted final approval of the class action settlement.[142] The settlement class included all United States persons whose credit, debit, or personal information was compromised as a result of the data breach. The settlement requires Target to pay $10 million into a fund to compensate class members for out-of-pocket losses and time lost, to pay $6.75 million in attorneys' fees, and to implement security measures.[143] The security measures include designating a Chief Information Security Officer, maintaining a written information security program, maintaining a process to monitor for and respond to information security events, and security training for employees.[144]

Three objectors and one non-class member appealed to the Eighth Circuit.[145] The district court questioned the merit of the appeals, categorizing some as "professional objectors" and another as a non-class member whose "appeal [was] frivolous."[146] In Miorelli v. Target Corp.,[147] the Eighth Circuit summarily dismissed the non-class member's appeal for lack of jurisdiction. Another objector's appeal, claiming that the magistrate judge's opinion stating that the settlement was fair and reasonable violated rules on expert testimony and usurped the role of the court, is still pending.[148]

The claims on behalf of a class of financial institution plaintiffs against Target developed further, but were also ultimately resolved through a class action settlement. On September 15, 2015, the court granted plaintiffs' motion to certify the class.[149] On May 12, 2016, Judge Magnuson granted final approval of the class action settlement.[150] The settlement requires Target to pay $20.25 million into a class escrow account on behalf of a class of entities that issued compromised payment cards, $19.1 million directly to fund MasterCard's Account Data Compromise program,[151] $100,000 total to five representative plaintiffs, and $19.9 million in attorneys' fees and expenses.[152] The settlement did not provide for any injunctive relief. Separately, Target reached a private settlement with Visa worth up to $67 million over claims that it failed to implement and maintain reasonable security procedures and practices to prevent the breach.[153]

c. Sony Pictures Entertainment

After a 2014 data breach, individual plaintiffs filed suits against Sony Pictures ("Sony") that were ultimately consolidated into a class action lawsuit on behalf of current and former employees whose information was compromised.[154] After the court denied in part Sony's motion to dismiss,[155] plaintiffs' moved to certify a class, but the parties agreed to a class action settlement agreement before the court ruled on plaintiffs' motion.[156] On April 6, 2016, Judge R. Gary Klausner in the Central District of California granted final approval of the class action settlement.[157] The settlement class comprises all current and former Sony employees and any individuals whose personally identifiable information was released in the breach.[158] The settlement requires Sony to provide two years of
identity theft protection services (in addition to the one year that Sony already provided following the breach), up to $2 million to reimburse expenses and time spent taking preventative measures to prevent identity theft, up to $2.5 million to reimburse loses from identity theft,[159] $33,000 in total service awards, and almost $2.6 million in attorneys' fees.[160] The settlement did not provide for any injunctive relief.

d. Historical Class-wide Settlements of Data Breach Claims

As reflected in the chart below, the class-wide data breach settlements this year are, on average, considerably more expensive than prior class-wide settlements. However, they consist of many of the same components as earlier settlements.

The chart reflects the relevant defendant, the date of final approval of the class-wide settlement, the data type involved in the data breach, the relief provided to the class as part of the settlement, and any fees and costs awarded to class counsel and service awards ordered for class representatives.

<table>
<thead>
<tr>
<th>Defendant</th>
<th>Approval</th>
<th>Data Type</th>
<th>Relief to the Class</th>
<th>Service Awards, Fees, &amp; Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home Depot (Consumer Class)[161]</td>
<td>August 23, 2016</td>
<td>Card Data</td>
<td>Up to $13 million for class claims; up to $6.5 million for 18 months of credit monitoring services; security practices changes</td>
<td>$1,000 for each representative plaintiff; $166,925 in costs; $7.536 million in fees</td>
</tr>
<tr>
<td>Target (Financial Institution Class)[162]</td>
<td>May 12, 2016</td>
<td>Card data</td>
<td>Up to $20.25 million for class claims;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$19.108 million to MasterCard</td>
<td>$20,000 for 5 representative plaintiffs; $2.109 million in costs; $17.8 million in fees</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reportedly up to $67 million for Visa's claims against Target[163]</td>
<td></td>
</tr>
<tr>
<td>Sony[164]</td>
<td>April 6, 2016</td>
<td>Login and Personal Information</td>
<td>Up to $2 million for preventative losses; up to $2.5 million for claims for identity theft losses; up to two years of credit monitoring services</td>
<td>$3,000 for each named plaintiff; $1,000 for each plaintiff who initially filed an action; $2.588 million in fees</td>
</tr>
<tr>
<td>Defendant</td>
<td>Approval</td>
<td>Data Type</td>
<td>Relief to the Class</td>
<td>Service Awards, Fees, &amp; Costs</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------</td>
</tr>
<tr>
<td>St. Joseph Health System[165]</td>
<td>February 3, 2016</td>
<td>Health Information</td>
<td>$7.5 million in cash payment; up to $3 million for class claims; one year of credit monitoring services (offered during remediation); security practice changes</td>
<td>$50,000 in incentive payments for class representatives; $7.45 million in fees and costs</td>
</tr>
<tr>
<td>Target (Consumer Class)[166]</td>
<td>November 17, 2015</td>
<td>Card Data</td>
<td>Up to $10 million for claims; security practice changes</td>
<td>$1,000 for three deposed plaintiffs; $500 for other plaintiffs; $6.75 million in fees</td>
</tr>
<tr>
<td>LinkedIn[167]</td>
<td>September 15, 2015</td>
<td>Login Information</td>
<td>Up to $1.25 million for claims; security practice changes</td>
<td>$5,000 for the named plaintiff; $26,609 in costs; $312,500 in fees</td>
</tr>
<tr>
<td>Adobe[168]</td>
<td>August 13, 2015</td>
<td>Login and Card Data</td>
<td>Security practice changes and audit</td>
<td>$5,000 to each individual plaintiff; $1.18 million in fees</td>
</tr>
<tr>
<td>Sony Gaming Networks[169]</td>
<td>May 4, 2015</td>
<td>Card Data and Personal Information</td>
<td>Up to $1 million for identity theft losses; benefit options including free games and themes or month subscription, unused wallet credits, virtual currency; some small cash payments</td>
<td>$2.75 million in fees</td>
</tr>
<tr>
<td>AvMed[170]</td>
<td>February 28, 2014</td>
<td>Personal Information</td>
<td>Up to $3 million; security practice changes</td>
<td>$5,000 for each representative plaintiff; $750,000 in fees</td>
</tr>
<tr>
<td>Defendant</td>
<td>Approval</td>
<td>Data Type</td>
<td>Relief to the Class</td>
<td>Service Awards, Fees, &amp; Costs</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>Purchasing Power (Winn-Dixie)[171]</td>
<td>October 4, 2013</td>
<td>Personal Information</td>
<td>Up to $225,000 for class claims; up to one year of credit monitoring services; security practice changes</td>
<td>$3,500 for representative plaintiff; $200,000 in fees</td>
</tr>
<tr>
<td>CBR Systems[172]</td>
<td>July 24, 2013</td>
<td>Health Information</td>
<td>Up to $500,000 for claims for expenses; up to $2 million for class claims for identity theft; two years of credit monitoring services; security practice changes</td>
<td>$5,000 for representative plaintiff; $14,064 in costs; $585,936 in fees</td>
</tr>
<tr>
<td>Michaels Stores (Pin Pad Litig.)[173]</td>
<td>April 17, 2013</td>
<td>Card Data</td>
<td>Up to $800,000 for class claims; up to two years of credit monitoring services; security practice changes</td>
<td>$2,500 for each representative plaintiff; $55,565 in costs; $1.2 million in fees</td>
</tr>
<tr>
<td>Heartland Payment Systems[174]</td>
<td>March 20, 2012</td>
<td>Card Data</td>
<td>Up to $2.4 million for class claims; security practice changes</td>
<td>$35,000 in costs; $606,193 in fees</td>
</tr>
<tr>
<td>Countrywide[175]</td>
<td>August 23, 2010</td>
<td>Personal and Financial Information</td>
<td>Up to $5 million for claims for identity theft; up to $1.5 million for claims for expenses; two years of credit monitoring services</td>
<td>$500 for each representative plaintiff; $250 for each named plaintiff; $100,000 in costs; $3.5 million in fees</td>
</tr>
<tr>
<td>Dep't of Veterans Affairs[176]</td>
<td>September 23, 2009</td>
<td>Personal Information</td>
<td>Up to $20 million for class claims</td>
<td>$18,000 for representative plaintiffs; $157,076 in costs; $3.6 million in fees</td>
</tr>
</tbody>
</table>
3. Shareholder Derivative Suits

In recent years, shareholders have sought to pursue derivative lawsuits against corporate directors and officers for breach of fiduciary duty in overseeing corporate security in connection with data breaches. In 2014 and 2015, shareholders brought three such high-profile derivative lawsuits on behalf of Wyndham Worldwide, Target, and Home Depot. The Wyndham suit was dismissed in October 2014, after the district court found that the board's actions were protected under the business judgment rule.[179] This year, the Target and Home Depot cases were similarly dismissed.

Target. After shareholders filed a derivative lawsuit in February 2014 in district court in Minnesota, Target's board of directors established a special litigation committee ("SLC") to investigate the claims at issue. On July 7, 2016, the court granted the SLC's unopposed motion to dismiss in reliance on a report issued by the SLC, which concluded that "it [was] not in Target's best interest to pursue" legal recourse against Target's directors and officers.[180]

The Home Depot. Home Depot shareholders filed a derivative lawsuit in September 2015 in district court in Georgia. On November 30, 2016, the court dismissed the action on grounds that shareholders failed to either demand that the board take action or demonstrate that such a demand would have been futile.[181] Since the Home Depot plaintiffs made no demand prior to filing suit, the court turned to the issue of demand futility.[182] To demonstrate demand futility under Delaware law, a plaintiff must plead particularized facts that establish reasonable doubt regarding the ability and willingness of the
board to evaluate a demand in a disinterested manner.[183] With regard to plaintiffs' primary claim for breach of the duty of loyalty, the court found that "[w]hen added to the general demand futility standard, the Plaintiffs essentially need to show with particularized facts beyond a reasonable doubt that a majority of the Board faced substantial liability because it consciously failed to act in the face of a known duty to act."[184] The court concluded that plaintiffs' allegations that the board violated this duty by disbanding Home Depot's infrastructure committee and moving too slowly in addressing the security breach were insufficient to overcome this "incredibly high hurdle."[185] After arriving at a similar conclusion for the claims for corporate waste[186] and violations of Section 14(a) of the Securities Exchange Act,[187] the court held that plaintiffs' failure to make a pre-suit demand was not excused, dismissed the case with prejudice, and permitted defendants to recover costs.[188]

While a number of additional high-profile data breaches have been announced since the filing of the Wyndham, Target, and Home Depot lawsuits,[189] shareholder derivative litigation has not kept pace with consumer class action litigation. This is likely because shareholders have come to recognize that there are substantial obstacles to proceeding on such claims—as amply demonstrated by the dismissal of the Wyndham, Target, and Home Depot suits at the pleadings stage. However, derivative lawsuits remain a concern for companies in this context. As noted above, in December, a Wendy's shareholder brought a derivative suit for breach of fiduciary duty, alleging that the board of directors and executive officers failed to protect Wendy's payment system in connection with a breach.[190] Companies should consider the risks of such lawsuits in connection with a company's implementation of any data breach response plan.

C. Interceptions and Eavesdropping

1. Email Scanning

As in past years, 2016 saw major developments in a number of ongoing class action lawsuits alleging that major Silicon Valley technology companies violated state and federal laws by scanning user emails and messages for targeting advertising and other business purposes. Companies operating electronic communications services should continue to monitor these suits, as they are often massive in scope, concern proposed classes including all or many users of a particular service, are predicated on alleged privacy violations that many perceive to be standard industry practices, and address the disclosures that satisfy consent to information collection and use.

Corley v. Google Inc. Further, one novel strategy employed in 2016 was an attempt by four plaintiffs who were members of the putative class in In re Google Inc. Gmail Litigation (where certification was denied) to sue as individuals. Specifically, in Corley v. Google, Inc., four UC Berkeley students sued Google for violations of the Electronic Communications Privacy Act ("ECPA"), alleging that Google scanned their college email accounts for Google's own commercial purposes and without users' consent.[191] The four plaintiffs then sought to join nearly 900 plaintiffs to the action, but Judge Koh rejected this end-run around certification, noting that whether "[individuals] have consented to the alleged interceptions has been central to this case" and that "both express and implied consent are questions of fact."[192] Judge Koh held that mass joinder was not appropriate because the claims did not "arise[] out of the same transaction, occurrence, or series of transactions or
Therefore, Judge Koh held that if the plaintiffs wished to proceed, they were required to do so via 876 separately filed complaints filed within 45 days. On October 3, 2016, the plaintiffs notified the court that all but two plaintiffs settled their claims with Google and moved to dismiss the claims with prejudice. On October 18, 2016, Judge Koh dismissed the claims of all plaintiffs with prejudice, including the two plaintiffs that failed to settle because they did not file an individual complaint within 45 days of the court's order.

*Matera v. Google Inc.* Another case to watch is *Matera v. Google, Inc.*, a case filed in September 2015 related to *In re Google Inc. Gmail Litigation*. There, the plaintiffs allege that Google's purported practice of collecting information on non-users violates CIPA and ECPA and, unlike in *In re Google Inc. Gmail Litigation*, seek both damages and injunctive relief. Because those allegations are related to those in *In re Google Inc. Gmail Litigation*, the court assigned the case to Judge Koh. On August 12, 2016, Judge Koh denied Google's motion to dismiss as to the merits of plaintiffs' claims. Specifically, with respect to the Wiretap Act, the court rejected Google's "ordinary course of business" argument and denied a motion to certify an interlocutory appeal to the Ninth Circuit on the same issue. Likewise, Judge Koh rejected Google's arguments that it should decline supplemental jurisdiction over the CIPA claim and that section 631 of CIPA does not apply to email communications.

On September 23, 2016, after lifting a stay pending the outcome of the Supreme Court's *Spokeo* decision, discussed herein *supra* Section II.A., Judge Koh granted in part and denied in part Google's motion to dismiss based on lack of standing. Most significantly, Judge Koh concluded that based on "the historical practice of courts recognizing that the unauthorized interception of communication constitutes cognizable injury" and "the judgment of Congress and the California Legislature [that] alleged violations of . . . the Wiretap Act and CIPA constitute injury in fact," the plaintiffs' complaint survived *Spokeo*. However, Judge Koh also held that plaintiffs lacked standing to enjoin Google from engaging in the alleged "intercepting and scanning," which Google confirmed it had ceased. On November 28, 2016, the parties in *Matera* requested a stay of the proceedings and announced that they had successfully mediated a resolution of the case and finalized a settlement agreement. The same day, Judge Koh granted the stay and ordered the plaintiffs to file a motion for preliminary approval of class action settlement by December 28, 2016. On December 13, 2016, plaintiffs outlined a settlement consisting of $2.2 million in attorneys' fees and $0 to members of the class. The settlement also allocated $2,000 for each of the two lead plaintiffs and $123,500 for the work of the settlement administrator. Google agreed to change certain technical aspects of its email processing including "eliminat[ing] any processing of email content that it applies prior to the point when the Gmail user can retrieve the email in his or her mailbox." Although class members will receive no monetary award under the terms of the proposed settlement, the release "extends solely to claims for declaratory, injunctive and non-monetary equitable relief." Other than the named representatives, no settlement class member will release any claim for monetary damages under CIPA or ECPA. A hearing on the proposed settlement is scheduled for March 9, 2017.

2. **Call Recording**

In recent years, there has been a flurry of lawsuits against businesses for recording customer phone calls without the requisite consent. The recording of telephone conversations is governed by a patchwork of
federal and state law. The federal Wiretap Act and most states allow such recordings as long as one party to the conversation--including the one doing the recording--consents to the recording.[209] Eleven states arguably require the consent of all parties to the call.[210] One of these all-party consent states is California, which has become the locus of call recording litigation.

During 2016, courts continued to clarify the contours of California's call recording laws, found in the California Invasion of Privacy Act ("CIPA"), California Penal Code § 630, et seq.[211] The statute now has potentially nationwide reach, as courts reinforced a 2011 Northern District of California holding that non-California plaintiffs can assert claims against a California defendant where the alleged violations occurred in California.[212]

On the class certification front, the court in Saulsberry v. Meridian Financial Services, Inc., continued the trend in declining to certify section 632 classes because "a consumer's objectively reasonable expectation regarding the confidentiality of the call depends on a great variety of individual circumstances, such as the individual's knowledge of the defendant's recording practices, or prior consent to the recording of the calls."[213] Section 632.7 does not include the "confidential communication" requirement that section 632 does, and thus there may be greater opportunities for class-wide litigation under section 632.7.[214]

However, section 632.7 is not without its own potential limitations. The court in Carrese v. Yes Online Inc. added its weight to an intra-circuit split among district courts over whether section 632.7 can only be enforced against third parties who "intercept or receive" the communication without consent.[215] Most district courts in the Ninth Circuit, including the court in Carrese, "have found section 632.7 applies both to parties of a communication as well as third parties."[216]

Another issue addressed in 2016 related to whether the statutory award of $5,000 was intended as a per violation award, or a per action award.[217] The court in Granina v. Eddie Bauer agreed with the weight of authority that damages are per violation, although it "strongly support[ed] an appellate decision clarifying this issue."[218]

Several notable settlements were reached in 2016 over alleged CIPA violations. For example, Wyndham International Inc. agreed in October to pay $7.3 million to settle class allegations that the hotel chain recorded customers' calls to the hotel's toll-free reservations hotline without notice that their calls would be recorded and without consent.[219] And, in August, HSBC Card Services Inc. agreed to a $13 million deal to settle three consolidated class actions alleging unlawful recording of debt-collection calls.[220]

In the public arena, the California Attorney General continued to aggressively enforce CIPA. In 2015, the California AG reached a settlement with Houzz Inc., an online platform for home remodeling and design, for alleged violations of CIPA.[221] It required Houzz to appoint a Chief Privacy Officer to oversee compliance.[222] In 2016 the AG reached a similar deal with a large financial institution, which required not only payment of penalties, but also a compliance program and designation of an individual to serve in a compliance oversight capacity.[223]

Another active area of call recording litigation involves the recording of inmate phone calls. In Romero v. Securus Technologies, Inc., two former inmates and a criminal defense attorney, all of whom used
Securus's telephone systems to make calls from California correctional facilities, sued Securus for secretly recording multiple attorney-client calls.[224] The plaintiffs alleged violations of section 636 of CIPA,[225] which prohibits recording, without all-party consent, conversations between "a person who is in physical custody of a law enforcement officer . . . or who is on the property of a law enforcement agency or other public agency, and the person's attorney."[226] On a motion to dismiss, the court dismissed various common law claims in the complaint but allowed the CIPA and unfair competition claims to advance.[227] In addition, the court drew a distinction between section 636, which does not require a showing that the communications were confidential, and section 632, which does, in rejecting the defendant's motion to strike the class allegations due to the supposed individualized inquiry that a section 636 confidentiality analysis would entail.[228] Plaintiffs filed an amended complaint on November 7, and Securus moved to dismiss on November 25.[229]

A similar suit against Securus for recording attorney-client conversations and disclosing those conversations to prosecutors, in violation of the federal and Texas wiretap acts, was settled in March 2016.[230] As part of the settlement, attorneys will be entitled to register their phone numbers on a "do not record" list that will maintain the confidentiality of their communications with clients who are in jail.[231] Securus was also recently sued by former NFL player Aaron Hernandez, who is serving a sentence for murder and facing trial on separate charges in Massachusetts. In a federal complaint, Hernandez alleged, among other causes of action, that Securus violated Massachusetts privacy law when recordings it had made of Hernandez's jailhouse phone calls were placed on an "unsecure electronic database" that was subsequently breached.[232]

3. Other "Interceptions"

Emails and telephone calls are not the only information that can be intercepted, and plaintiffs are increasingly bringing lawsuits based on the interception and collection of other types of information. These types of actions are now winding their way through the court system, and several saw important developments in 2016. The Sixth Circuit also issued a ruling that could expand liability to include actions taken by users of a defendant's products.

**Raney v. Twitter, Inc.** In *Raney v. Twitter, Inc.*, plaintiff brought a putative class action against Twitter alleging that Twitter read its users' direct messages and replaced hyperlinks within the messages with its own custom links.[233] Plaintiff brought claims for violations of ECPA and state privacy statutes, and invasion of privacy. After Twitter filed a motion to dismiss explaining that it scans the links stored on its servers, but not while the messages are in "transit" (as required for liability under the ECPA), plaintiff voluntarily dismissed the suit.[234]

**Audio Beacon Cases.** In the latest twist on theories of "interception," several professional sports teams have been hit with lawsuits alleging that their team apps illegally spy on fans using "beacon" technology. In a lawsuit against the NBA's Golden State Warriors, plaintiffs alleged that the team's app employs "beacons," which allow users to receive promotions on their phones based on their location whenever the phone detects an audio signal emitted by a "beacon."[235] According to plaintiffs, this beacon technology records and monitors conversations in violation of the Wiretap Act, even when a user's phone is turned off, a claim that the Warriors have flatly denied.[236] In October, the NFL's
Indianapolis Colts were hit with a similar lawsuit, again premised on beacon technology that allegedly listens to fans' conversations when activated.[237]

**Luis v. Zang.** In a consequential decision for companies that create software or hardware that can be used to intercept communications, the Sixth Circuit held that a software manufacturer could be directly liable for interceptions accomplished with the company's software under federal and state wiretap laws. In *Luis v. Zang*, a suspicious husband surreptitiously installed a product known as "WebWatcher" to monitor his wife's online communications with another man.[238] The software allegedly intercepted "all PC activity including emails, IMs, websites visited, web searches, Facebook/MySpace activity, and anything typed in real time."[239] The Sixth Circuit first held that the maker of WebWatcher--Awareness Technologies, Inc.--could be liable for manufacturing, marketing, selling, or operating a wiretapping device in violation of 18 U.S.C. § 2512(1)(b), which creates liability for manufacturers of devices that are "primarily useful for the purpose of... surreptitious interception."[240] But the court did not stop there. Whereas the district court had held that only the disgruntled husband could be liable for the interceptions themselves, the Sixth Circuit reversed and held that Awareness could also be liable for the "interceptions," because the software "automatically acquires and transmits communications to servers that Awareness owns and maintains" without "any active input from the user."[241] In so holding, the Sixth Circuit may have opened the door to more expansive liability for companies whose products are used for illegal interceptions.

**D. Telephone Consumer Protection Act**

The TCPA[242] continued to be a popular statute for the plaintiff's bar in 2016. Part of this is likely attributable to the statute's provision of statutory damages in the range of $500 to $1,500 per violation. But another possible factor is a plaintiff-friendly 2015 FCC omnibus order that, among other things, defined an autodialer to include any equipment with the "potential ability" to store or produce telephone numbers to be called and to call those numbers--not solely equipment with the *current* capability to do this.[243] The omnibus order also made clear that separating the equipment that stores the number from the equipment dialing the number, even in different organizations, may not suffice to avoid the TCPA's applicability.[244] The omnibus order also changed the way that a consumer can revoke consent: now, not only may "a called party . . . revoke consent at any time and through any reasonable means" but "[a] caller may not limit the manner in which revocation [of consent] may occur."[245]

The FCC's 2015 order drew significant fire, and the D.C. Circuit Court of Appeals is currently weighing a closely-watched appeal that will determine the order's continued viability. While the D.C. Circuit has yet to rule on the appeal, it heard oral argument on October 19, 2016.[246] During oral argument the petitioners--including debt collector trade group ACA International, the company Salesforce.com, and the U.S. Chamber of Commerce--argued that the FCC's vague definition of what amounts to an autodialer encompasses all modern smartphones, which theoretically have the ability to download an application that allows for autodialer-like features. Petitioners' argument on this point may have had traction, as during the hearing the three judge panel pushed back and questioned whether the purported expansion of the autodialer definition went beyond the FCC's power. The FCC denied this, arguing that
the statute itself is ambiguous at best, and that the FCC has the authority to interpret the language in a reasonable manner.

With this appeal still pending, litigants and district courts have grappled with how to proceed in TCPA litigation. Some courts have elected to stay TCPA cases pending the D.C. Circuit's decision on the grounds that defendants seeking a stay could face difficulties during discovery because, for instance, of the unclear distinction between "potential" and "theoretical" capacity in the definition of an autodialer. But other courts have elected to proceed, reasoning that the delay imposed by a stay could turn out to be indefinite if the D.C. Circuit's ruling is ultimately appealed to the U.S. Supreme Court.[247]

Courts in 2016 also navigated the impact that the Supreme Court's ruling regarding standing in *Spokeo, Inc. v. Robins* has in TCPA actions.[248] In *Spokeo*, the Court held that a plaintiff cannot "automatically satisfy[...y] the injury-in-fact requirement whenever a statute grants a person a statutory right and purports to authorize that person to sue to vindicate that right."[249] For the most part, post-*Spokeo* TCPA decisions seem willing to allow plaintiffs to move forward even with a showing of minimal injury. For instance, courts have found that allegations regarding a plaintiff's time spent answering or addressing robocalls, a plaintiff's telephone line being rendered unavailable as a result of a robocall, or even the depletion of a cellular telephone's battery, were all sufficient to meet *Spokeo's* concrete injury requirement.[250] On the other hand, courts have dismissed TCPA cases for lack of standing where the alleged injury could not be connected to a particular violation of the TCPA.[251] In short, post-*Spokeo* plaintiffs must generally allege some minimal injury to themselves or their affected devices resulting from a defendant's alleged TCPA violations.

This year also saw what may be the largest TCPA settlement in history. The parties in *Aranda v. Caribbean Cruise Line, Inc.* agreed to a settlement with a common fund of at least $56 million—and up to $76 million—to end litigation that has been pending since 2012.[252] The *Aranda* settlement would allocate a maximum of $500 per call to class members (recipients of the defendant's 900,000 illegal robocalls), although the ultimate amount will be determined based on the size of the settlement fund and the number of claimants.[253] The settlement was reached after the court granted class certification in August 2014, denied a motion to decertify the class, and granted partial summary judgment to plaintiffs, holding that the calls at issue violated the TCPA. The *Aranda* court granted preliminary approval of the settlement in October 2016; it is expected to decide on final approval in February 2017.[254]

The future of TCPA litigation in 2017 likely will be impacted by the D.C. Circuit's decision as to the omnibus order, and marked by courts continuing to determine the applicability and limits of *Spokeo*. Finally, with a new administration change, and new FCC Chair, the contours of TCPA litigation also are likely to be marked by any new omnibus orders the FCC may issue interpreting the TCPA.

**E. Video Privacy Protection Act**

In 2016, courts continued to grapple with the contours of the VPPA,[255] which passed in 1988 in response to a D.C. newspaper's attempt to embarrass Judge Robert Bork during his Supreme Court nomination hearings by publishing his video store rental records.[256] As in recent years, courts in 2016 often came to differing conclusions when applying the VPPA to modern technologies that did not exist...
at the law's inception, and plaintiffs continued to exploit these ambiguities because the stakes of violating the VPPA are high. Indeed, the law provides a minimum $2,500 per-person in statutory damages (as well as attorneys' fees) when "video tape service providers" "knowingly" disclose "personally identifiable information concerning any consumer" to third parties, with certain limited exceptions.[257]

Courts also continued to disagree about who is a "subscriber," and thus a "consumer," under the VPPA, with at least one appeals court endorsing a much broader interpretation of the term than in years past. The VPPA defines a "consumer" as "any renter, purchaser or subscriber of goods or services from a video tape service provider."[258] Last year, in Ellis v. Cartoon Network, the Eleventh Circuit held that "downloading an app for free and using it to view content at no cost is not enough to make a user of the app a 'subscriber'" under the VPPA.[259] The appeals court found that there was no "ongoing commitment or relationship between the user and the entity" sufficient to make the plaintiff a "subscriber" because the plaintiff had not established a Cartoon Network account or profile, provided any "personally identifiable information" to Cartoon Network, paid for the app, or signed up for any periodic "services or transmissions" or access to exclusive content. Id.[260] Rejecting the Eleventh Circuit's interpretation of the term, the First Circuit held in Yershov v. Gannett Satellite Information Network Inc. in September 2016 that an individual who merely downloaded a free mobile app and watched free video clips was a "subscriber" under the VPPA.[261] Following the reasoning in Ellis, the lower court dismissed the suit in May 2015 on the grounds that the plaintiff's mere use of a free mobile app did not make him a "subscriber," as he had not paid any money, registered any information, or received a delivery to access the app. On appeal, the First Circuit held that a monetary payment is not a necessary condition to be a "subscriber" under the VPPA, further reasoning that in downloading the app and providing "personally identifiable information" in the form of device ID and GPS coordinates, the plaintiff was "not free of a commitment to provide consideration in the form of that information."[262]

Further, in affirming the lower court's finding that device ID and GPS coordinates constitute "personally identifiable information" under the VPPA, the First Circuit also departed from several other courts in concluding that "personally identifiable information" is not simply limited to "information that explicitly names a person."[263] The First Circuit remanded the case, and the district court was charged with deciding whether the plaintiff had sufficiently pled his injuries to continue his case under the Supreme Court's recent Spokeo ruling. The district court ultimately denied the defendant's motion to dismiss, ruling that "the intangible harm allegedly suffered by [the plaintiff] from Gannett's alleged disclosure of his [personally identifiable information] is a concrete injury in fact."[264]

Courts addressed the First Circuit's reasoning in Gannett in subsequent VPPA cases. For example, in June 2016, the Third Circuit held that the VPPA's prohibition on the disclosure of "personally identifiable information" applies only to the kind of information that would readily permit an ordinary person to identify a specific individual's video-watching behavior--and accordingly, "static digital identifiers" such as IP addresses, and browser and operating system settings, are outside the purview of "personally identifiable information" under the VPPA.[265] However, the appeals court insisted that its decision does not create a split with the definition endorsed in Gannett and stressed that it intended to "articulate a more general framework" rather than establish a sweeping, broadly applicable rule "given the rapid pace of technological change in our digital era."[266] However, the Third Circuit explicitly punted the question of what other kinds of disclosures can trigger liability under the statute to "another day" and
cautioned that "companies in the business of streaming digital video are well advised to think carefully about customer notice and consent" while such issues get sorted out.[267]

The Eleventh Circuit now has an opportunity to embrace or reject the reasoning in Gannett in a VPPA appeal that is currently pending. In Perry v. Cable News Network, a federal district court in Atlanta embraced Ellis's reasoning and dismissed the action on the grounds that the plaintiff was not a "consumer" under the VPPA where there was "no indication that he had any ongoing commitment or relationship with defendants, such that he could not simply delete the CNN App without consequences."[268] The defendant maintains that Ellis mandates dismissal because the data it allegedly sent out--a random device identifier--was not "personally identifying," and that the Gannett case was an "outlier lacking any meaningful limiting principle."[269] This is an appeal to watch.

Additionally, in the wake of the Supreme Court's Spokeo decision, discussed above in Section II.A.1., courts have continued to address plaintiffs' standing to bring claims under the VPPA and its state-law analogs. For example, a New York federal judge rejected Conde Nast's motion to dismiss a suit brought under Michigan's state law analog to the VPPA (the Michigan Preservation of Personal Privacy Act) for lack of Article III standing. As in the Gannett case discussed above, the court noted that Spokeo provides that "'concrete' doesn't necessarily mean 'tangible," and that "intangible injuries can . . . be concrete," and it rejected the argument that the plaintiff pled only a "harmless procedural violation" in alleging that her personal information was unlawfully disclosed and used as a result of Conde Nast's practices.[270] Critically, the court cited several recent VPPA cases in noting that "all courts to consider the question, including this one, have concluded--both pre-and post-Spokeo--that consumers alleging that a defendant violated the VPPA by 'knowingly disclos[ing] their [personally identifiable information] to a third party without their consent have satisfied the concreteness requirement for Article III standing.'"[271] The court also noted that "post-Spokeo VPPA decisions recognized that Congress may elevate an otherwise non-actionable invasion of privacy into a concrete, legally cognizable injury," and that the harms contemplated by both the VPPA and the Michigan Preservation of Personal Privacy Act "have close ties to those recognized by the common law tort of invasion of privacy."[272] Accordingly, the court found that the plaintiff sufficiently alleged a "concrete, if hard to measure, intrusion on protected privacy interests."[273]

F. California's Song-Beverly Credit Card Act and Point-of-Service Data Collection

In 2016, plaintiffs and defendants alike used novel strategies in contending with recent court decisions that limited the scope of California's Song-Beverly Credit Card Act of 1971 ("Song-Beverly"),[274] which prohibits merchants from requesting or requiring a customer's personal identification information as a condition of accepting a credit card payment. Recent decisions narrowing Song-Beverly's reach have been deeply felt because, limited statutory exceptions notwithstanding, the prohibitory language of the law sweeps broadly, and those found in violation face civil penalties of up to $250 for the first violation and up to $1,000 for subsequent violations.[275]

In 2016, some plaintiffs pressed courts to focus not on the precise timing of a merchant's request for a customer's personal information, but to instead look more broadly at whether the customer reasonably understood that the request was optional in determining whether it was a condition of the transaction
under Song-Beverly. Such a test would arguably circumvent a controlling, bright-line test established in 2015 by a California appellate court that provides that a brick-and-mortar retailer does not violate Song-Beverly by requesting email addresses after credit card transactions are concluded because customers cannot reasonably believe that providing such information is a "condition of acceptance of the credit card."[276] For example, in October 2016, a plaintiff filed a notice of appeal of a trial victory granted by a San Diego County judge to Urban Outfitters in August 2016 on the basis that the collected ZIP code information was not a condition of payment under Song-Beverly since cashiers were only prompted to request such information from customers after the credit card had been swiped, approved, and signed for.[277] The plaintiff's counsel has asserted that the plaintiff intends to argue on appeal that a Song-Beverly violation should only be found if any reasonable consumer could believe the information is required as a condition of payment.[278] Relatedly, in a suit that was dismissed after the parties reached a non-public settlement, the plaintiff's complaint focused not on the timing of the request for personal information, but on allegations that he "reasonably believ[ed] that he was required to provide the requested information to complete the transaction."[279] Whether courts will embrace this standard favored by plaintiffs remains to be seen.

Both plaintiffs and defendants in 2016 also invoked the Supreme Court's Spokeo v. Robins decision in challenging the standing of plaintiffs alleging Song-Beverly violations. Most notably, in October 2016, in Fraser v. Wal-Mart Stores, Inc., the U.S. District Court for the Eastern District of California rejected Wal-Mart's argument that Spokeo required dismissal of a class action accusing the retailer of unlawfully collecting shoppers' ZIP codes because the suit alleges only a single procedural violation of Song-Beverly, holding instead that consumers' alleged exposure to "undesired marketing contact" and the real risk of identity theft constituted sufficient concrete harm to confer Article III standing.[280] In another interesting case, a plaintiff in Medellin v. Ikea U.S. West Incorp. invoked Spokeo and its progeny in her bid for the Ninth Circuit to dismiss her Song-Beverly appeal for lack of subject-matter jurisdiction and direct the district court to remand the matter to state court.[281] There, the plaintiff argued that her case is similar to the D.C. Circuit's recent decision in Hancock v. Urban Outfitters, Inc.,[282] which acknowledged the plaintiff lacked Article III standing under Spokeo because he merely alleged a "bare violation of the law without more," and the lower court never had jurisdiction to hear the suit in the first place. In Medellin, Ikea countered that the plaintiff alleged the concrete harm necessary to proceed with her appeal by asserting that collecting customers' ZIP codes subjected them to increased risk of identity theft, fraud, and invasions of privacy. On January 13, 2017, the Ninth Circuit--in an unpublished order--vacated the district court's judgment and remanded with instructions that the district court dismiss the action without prejudice for lack of standing, on the grounds that a plaintiff cannot "allege a bare procedural violation, divorced from any concrete harm, and satisfy the injury-in-fact requirement of Article III."[283]

Finally, the full implications of the Ninth Circuit's landmark December 2015 ruling in Big 5 Sporting Goods remains to be seen. In Big 5, consumers filed eleven class action suits alleging that Big 5 infringed on privacy rights by requesting, recording, and publishing customer ZIP codes during credit card transactions in violation of Song-Beverly.[284] Big 5 subsequently sued its insurers after they refused to provide it with a defense against the lawsuits.[285] The Ninth Circuit affirmed a grant of summary judgment to the insurers, and, like the Third Circuit three months earlier in OneBeacon America Insurance Co. v. Urban Outfitters Inc.[286] held that the "statutory violation" exclusions in general
liability insurance policies barred "personal and advertising injury" coverage for underlying allegations of unlawful ZIP code collection.\[287\] In other words, insurers have no duty to defend their insured against underlying claims that they infringed on privacy rights in violation of Song-Beverly. The Ninth Circuit further held that the policy exclusion precluded the duty to defend even where an underlying action alleges common law violations of invasion of privacy, since California does not recognize any common law or constitutional privacy causes of action for "garden variety ZIP Code cases like this."\[288\]

While courts did not address the effect of Big 5 on ZIP code coverage suits under Song-Beverly in 2016, the decision may play a key role in determining whether retailers may look to their insurers to cover costs stemming from defense of Song-Beverly actions. Already, at least one insurer has cited Big 5 in an appeal of a data privacy suit alleging a different statutory violation.\[289\]

G. Biometric Information Privacy Acts

Litigation centered on biometrics became increasingly common in 2016. This is in part because companies are increasingly relying on biometrics--distinctive physiological characteristics, such as fingerprint, hand or face geometry, retina scans, or voice patterns--to authenticate an individual's identity, or otherwise integrating them into everyday services such as photo-sharing websites. But because biometrics are biologically unique, they also present particular concerns. For instance, if a person's social security number is compromised, it can be changed. Biometrics cannot.

Recognizing this risk, some states have started regulating the collection, use, and storage of biometric information. Several states have introduced biometric legislation,\[290\] but only two states have enacted laws. Illinois has passed BIPA,\[291\] followed shortly thereafter by Texas's biometric protection law, the Texas Business and Commerce Code Section 503.001.\[292\] However, only Illinois' law has a private right of action. Key components of BIPA include its requirements that private companies obtain informed written consent prior to collecting a person's biometric identifier or information, and meet certain protection obligations and retention guidelines.\[293\] Significantly, BIPA also permits the recovery of damages of up to $5,000 per violation.\[294\]

BIPA has recently become a magnet for the plaintiffs' bar, with putative class actions filed against companies such as Facebook, Google, Snapchat, and Shutterfly based on those companies' uses of facial recognition technology.\[295\] In each of these suits, the crux of the allegations was the same: the defendant allegedly failed to obtain the plaintiff's informed consent before scanning the plaintiff's face after his or her image was uploaded to the company's photo-sharing platform, and failed to adhere to BIPA's requirements relating to the disclosure and destruction of biometric identifiers. The defendants in most of these cases have taken the position that, because the biometrics they collect come solely from photographs, BIPA does not apply.\[296\] This is because, although BIPA's statutory language states that scans of "hand or face geometry" are biometric identifiers, it also expressly provides that "photographs" are not.

One decision on this issue is a 2016 order denying a motion to dismiss in a putative class action pending against Facebook in the Northern District of California. In In re Facebook Biometric Information


Privacy Litigation, plaintiffs allege Facebook's photograph tagging suggestion feature, which uses facial recognition to analyze photos uploaded by users and "suggests" which of the user's Facebook friends is pictured, violates BIPA. In its motion to dismiss, Facebook argued that, because its facial recognition technology is used to analyze photographs uploaded to its service, after which data derived from those photographs is used to offer tagging suggestions, it falls within the exception to BIPA for photographs and information derived from them. U.S. District Court Judge Donato disagreed, interpreting the statute's use of the term "photographs" as meaning "paper prints of photographs, not digitized images stored as a computer file and uploaded to the Internet" (although he did not offer a statutory basis for this interpretation). It remains to be seen whether other courts will follow this interpretation of BIPA.

In addition, defendants in BIPA litigation may have new hope after the U.S. Supreme Court's May 16, 2016 decision in Spokeo, Inc. v. Robins, which further clarified the injury-in-fact component of the Article III standing requirements. In one of the only decisions applying Spokeo to BIPA to-date, McCullough v. Smarte Carte, Inc., the Northern District of Illinois recently dismissed a BIPA action for lack of standing, emphasizing the necessity of a "concrete and particularized injury" as articulated in Spokeo. The McCullough plaintiff alleged that a locker rental company violated BIPA by retaining her fingerprint without written consent. While acknowledging that the defendant technically violated BIPA, the court--relying on Spokeo--found the plaintiff failed to allege any harm that resulted from the violation and held that such a "bare procedural violation … cannot satisfy Article III standing."

The McCullough decision represents not only a deterrent to prospective BIPA plaintiffs going forward, but also a potential turning point in pending BIPA litigation. Indeed, Facebook recently filed a motion to dismiss for lack of subject matter jurisdiction based on Spokeo in In re Facebook Biometric Information Privacy Litigation, although the district court there has not yet issued a decision. Another case to watch on the issue of facial recognition technology and standing under Spokeo is Vigil et al. v. Take-Two Interactive Software, Inc., a class action filed in October 2015 in New York against video game company Take-Two Interactive over a face-scanning feature in its basketball games that allows players to create an in-game character with the player's likeness. Plaintiffs there contended that Take-Two never obtained their written consent before disseminating the information it gathered through the feature, but Take-Two countered that the lawsuit misuses BIPA to attack a feature that permits a player to create a cartoon-like in-game character that may or may not actually look like the player. Plaintiffs were permitted to submit a second amended complaint following the Spokeo ruling, and Take-Two filed a motion to dismiss arguing plaintiffs had not alleged harm sufficient to meet Spokeo's "concrete and particularized injury" requirement. The district court has yet to rule on Take-Two's motion to dismiss. A similar motion to dismiss based on Spokeo is pending in at least one other BIPA lawsuit. Once decided, these cases could help determine the course of future BIPA litigation.

H. Internet of Things and Device Hacking

The Internet of Things ("IoT") is continuously growing as traditionally "dumb" devices are transformed into connected and smart devices. No longer is IoT limited to smart phones and webcams; today, IoT also includes medical devices, routers, lighting, heating, and self-driving cars. Throughout 2016 there
was an increase in regulatory and private actions and additional guidance from regulators related to IoT corresponding with this growth.

**Routers, Cloud Storage, and Connected Cameras.** On February 23, 2016, ASUSTeK Computer Inc., a Taiwanese router manufacturer, agreed to a proposed consent order to resolve the FTC's probe into claims that security flaws in both the router and "cloud" storage services left users' personal information vulnerable to hackers and viruses.[307] The order required that ASUSTeK refrain from misrepresenting the security of its routers, establish and implement a comprehensive security program, provide consumers an opportunity to register for direct security notifications, and notify consumers directly of any software updates.[308] The FTC stated that this order was unique in its requirement that ASUSTeK offer consumers a way to register to receive security notices through direct communication, like email, text message, or push notification, therein providing a form of notification that goes beyond a posting on the manufacturer website that might go unnoticed by consumers.[309] Moreover, the registration for this type of notification must not be dependent upon or default to an agreement to receive non-security-related notifications, such as advertising.[310]

Just this month, on January 5, 2017, the FTC sued D-Link, a provider of wireless routers and IP-connected cameras, in the Northern District of California for violations of the FTC Act.[311] The FTC alleges that D-Link advertises its routers and cameras as containing "Advanced Network Security," but that flaws in D-Link's security allow hackers to easily access consumers' information and cameras.[312] The complaint further alleges that these security vulnerabilities put consumers at risk of harm by, for example, re-directing consumers "seeking a legitimate financial site to a spoofed website, where they would unwittingly provide the attacker with sensitive financial account information," by obtaining financial documents, such as tax returns, "stored on the router's attached storage device," or by using a compromised connected camera to monitor consumers' whereabouts.[313] The complaint against D-Link alleges one count of unfairness relating to D-Link's failure to secure consumer's information and five counts of misrepresentation relating to D-Link's advertising and statements that its routers and internet cameras are secure.[314] An initial case management statement is due by March 30, 2017.[315]

**Connected and Autonomous Automobiles.** In November 2015, in *Cahen v. Toyota Motor Corp.*, U.S. District Judge Orrick granted defendant car manufacturers Toyota, Ford, and General Motors' motions to dismiss a class action complaint alleging, among other claims, that the vehicles' computers were vulnerable to hacking and privacy violations related to their computer software.[316] Plaintiffs appealed to the Ninth Circuit, arguing that the district court erred in holding that plaintiffs failed to establish standing to assert their claims.[317] The parties completed briefing on November 9, 2016, but oral argument has not yet been scheduled.[318]

Another federal district case was filed after Chrysler and Harmon International Industries voluntarily recalled their vehicles because the vehicle computer system (uConnect) had design vulnerabilities that could allow hackers to take remote control of the vehicle's functions.[319] In *Flynn v. FCA US LLC*, plaintiffs alleged that these vulnerabilities violated the Magnuson-Moss Warranty Act and Michigan, Illinois, and Missouri state laws.[320] In September 2016, Chief Judge Reagan rejected plaintiffs' standing theory based on risk of harm or a fear of risk of harm that a future car hacking could injure or
kill them,[321] but accepted their theory that they overpaid for their vehicles because the vehicles were initially defective and the ongoing vulnerabilities have diminished their vehicles’ values.[322] The judge lifted the stay against the remaining two plaintiffs' claims on January 10, 2017 and set a new briefing schedule for the parties with defendants' motion to dismiss due by February 6, 2017.[323]

The potential for security breaches and privacy violations related to self-driving and other automobile software is a topic that has drawn and will continue to draw regulatory scrutiny. For example, in September 2016, the U.S. Department of Transportation released a Federal Automated Vehicles Policy that included guidelines on data collection and security.[324]

**Smart TVs.** Private actions against manufacturers of connected devices, while still in their nascent stages, have also brought to light the data privacy and security issues applicable to IoT. Smart TV manufacturer Vizio Inc. is in the midst of a multidistrict class action litigation defending against claims that the company violated customers' privacy rights by installing tracking software into its smart TVs that allowed it to collect viewing data and share it with third parties.[325] Plaintiffs allege that this data was then used to push targeted advertisements to the smart TVs, as well as to other connected devices that shared the same internet connection.[326] On April 11, 2016, the U.S. Judicial Panel on Multidistrict Litigation consolidated the 20 class action claims to the Central District of California.[327] Vizio filed a motion to dismiss, and the court heard oral argument on this motion on December 16, 2016.[328]

**Smart Toys.** Connected toys have also been the subject of private actions, including a recent proposed class action against ToyTalk, Inc. and Mattel, Inc. alleging that the toy Hello Barbie recorded and stored the voices of children without obtaining adequate consent in violation of the Children's Online Privacy Protection Act ("COPPA").[329] In this case, which was filed in California Superior Court on December 7, 2015, and subsequently removed to the Central District of California on March 29, 2016, the connected Barbie doll included a smartphone app that would allow the parents of the child to listen to, review, and delete recordings that the Barbie doll transmitted to ToyTalk's services.[330] The complaint alleged that while consent was obtained by the parents whose child owned the toy, the doll also captured the voices of other children whose parents had not consented to the use of the doll.[331] Though this case was ultimately voluntarily dismissed on July 22, 2016,[332] it raises unique issues that connected toys will encounter, such as compliance with COPPA.[333]

**Regulator Response.** As these private actions are just starting to test the boundaries of data privacy and security in IoT and clear precedents have yet to be developed, regulators have provided some guidance for best practices in relation to data security and privacy of connected devices. In January 2015, the FTC released a staff report summarizing the FTC's November 2013 workshop and providing staff recommendations related to IoT.[334] Notably, the report stated that FTC staff did not believe that IoT-specific legislation is needed at this time, but rather recommended that Congress should enact general data security legislation to strengthen the FTC's existing data privacy and security tools. In November 2016, the National Institute of Standards and Technology (NIST) released guidance on building security safeguards directly into connected devices and included technical standards and security principles that developers are advised to take into account during every phase of a product's development.[335] These issues have also garnered the attention of the California executive branch. In the wake of the October
21, 2016 Distributed Denial of Service (DDoS) attack on Dyn, an internet infrastructure company, that caused massive internet outages, former California Attorney General Kamala D. Harris acknowledged the unique security vulnerabilities of connected devices and urged consumers to change the passwords of their household connected devices.[336] In addition, the U.S. Food and Drug Administration issued guidance on December 28, 2016, outlining post-market recommendations for medical device manufacturers.[337] Specifically, these guidelines recommend implementing comprehensive cybersecurity risk management programs and responding in a timely fashion to identified vulnerabilities.[338]

I. Cybersecurity Insurance

In the face of the growing threat of, and costs associated with, cyberattacks, approximately one-third of U.S. companies have turned to insurance providers for protection.[339] State and local governments are also increasingly adding cybersecurity insurance to their policies. For example, Idaho recently acquired a $25 million policy following a breach of the state's fish and game data held by a third-party vendor.[340]

Because commercial insurance policies do not generally cover many cyberattacks, carriers have started offering standalone cybersecurity policies.[341] More than 70 insurance companies currently offer such policies; however, that group represents less than 2% of the insurance industry overall.[342] Thus, while there are certainly more cybersecurity insurance policies available now than in past years, cyber insurance is still very much a developing field, and securing affordable policies remains challenging for many businesses. And while "U.S. insurers are becoming more skilled at underwriting and pricing standalone cyber insurance policies,"[343] the limited publicly available actuarial data concerning the scale and financial impact of cyberattacks has resulted in significant variations in cybersecurity insurance premiums.[344] However, companies such as California-based CoverHound appear to be working to develop cyber insurance policies packaged with risk monitoring programs, in order to help small companies shore up their network protections.[345]

In this nascent market, litigation has arisen over the scope of cyber insurance coverage. For example, earlier this year, restaurant chain P.F. Chang's filed a lawsuit against its cybersecurity insurance provider for reimbursement of third-party costs associated with a data breach.[346] P.F. Chang's credit card processor, Bank of America Merchant Services ("BAMS"), incurred approximately $1.9 million in costs as a result of the breach--including costs associated with notifying cardholders, issuing and delivering new cards and account numbers, and covering fraudulent charges--and sought reimbursement from P.F. Chang's pursuant to a contract between the parties.[347] P.F. Chang's reimbursed BAMS, and then sought repayment through its insurance policy with Federal Insurance. Federal Insurance refused, claiming that P.F. Chang's payment to BAMS was not covered. The court agreed, concluding that P.F. Chang's policy was not meant to cover certain fraud recovery costs, such as its costs of reimbursement to BAMS, and that the policy excluded costs associated with P.F. Chang's contractual obligations to third parties.[348] The court also found that P.F. Chang's had no reasonable expectation that its policy would cover such costs, and that P.F. Chang's could have specifically negotiated for such coverage.[349]
The consistent flow of new entrants into the cyber insurance market means that there is little standardization in policy offerings, making it important for organizations to work closely with carriers to draft policies that best suit their needs. Particular attention should be paid to what is excluded from coverage, such as acts of terrorism, and to exclusions based on location of data storage and type of data affected. Additionally, as the P.F. Chang's case illustrates, companies should consider whether and to what extent payments to third parties are excluded from coverage under their own cybersecurity insurance policies. Despite the many factors to consider when crafting a policy, industry analysts and government agencies agree that cyber insurance is an important and beneficial tool to pursue in light of the current cyber landscape.

II. U.S. Government Regulation of Privacy and Data Security

A. Enforcement and Guidance

1. Federal Trade Commission ("FTC")
   
a. Data Security Enforcement

With its regulatory authority affirmed by the Third Circuit's decision in F.T.C. v. Wyndham Worldwide Corp., the FTC continued to bring enforcement actions against corporations for faulty data security practices throughout 2016.

LabMD. In July, the Federal Trade Commission found that the now-defunct company's data security practices were "unfair" and thus in violation of Section 5 of the FTC Act. LabMD had allegedly failed to take basic precautions to protect sensitive consumer information, and this resulted in billing information for 9,300 consumers becoming accessible on a peer-to-peer network and other personal information for at least 500 consumers ending up in the hands of identity thieves. In 2015, an administrative law judge ("ALJ") dismissed the FTC's charges for failure to demonstrate that LabMD's conduct created a "probability" or likelihood of harm. However, the final decision by the full Commission reversed the ALJ's ruling and held that LabMD failed to reasonably protect its customers' personal information from data breaches. The order requires LabMD to establish a comprehensive information security program to safeguard personal consumer information in its possession. LabMD must also obtain periodic independent assessments of its data security practices as well as notify consumers whose personal information was stolen in the data breach. However, the enforcement order has been stayed by the Eleventh Circuit pending LabMD's appeal. The panel found that LabMD was in a uniquely distressed position given that it is now defunct, and held that the costs of compliance would constitute an irreparable harm. The LabMD case will continue to be closely watched in 2017, as in its appellate briefs LabMD argues that the FTC does not have the broad authority to regulate cybersecurity practices. A decision by the Eleventh Circuit in LabMD's favor could create a circuit split on the issue.

ASUS. The FTC also resolved a data security enforcement action against computer hardware manufacturer ASUS. The FTC had charged ASUS with failing to take reasonable steps to secure the software on its routers, claiming that hackers had exploited vulnerabilities to access more than 12,900 consumers' connected storage devices. In July 2016, the FTC entered into a final consent decree...
that called for ASUS to establish "a comprehensive security program subject to independent audits for the next 20 years."[362] ASUS was also required to notify consumers about software updates and how users could protect themselves from security flaws, as well as provide an option for users to register for direct security notices.[363] Besides the security flaws themselves, the FTC also appeared to take issue with ASUS's marketing representations. The FTC noted that the company claimed its products "could protect computers from any unauthorized access, hacking, and virus attacks."[364] Companies should therefore carefully consider what their advertisements and marketing materials say about the data security of their products or services.

Mobile Devices. In May 2016, the FTC launched an inquiry to investigate vulnerabilities in mobile devices.[365] The regulator issued orders to eight major mobile device manufacturers, including Google and Samsung, requiring them to provide information about how the companies address security flaws in their products.[366] At the same time, the Federal Communications Commission ran a separate, parallel inquiry into common carriers' policies regarding mobile device security updates.[367] Among the selected common carriers were AT&T, T-Mobile, and Verizon.[368]

b. Privacy Enforcement

InMobi. The FTC also continued its practice of regulating use of customers' personal information. In June 2016, it entered into a consent decree with InMobi, a mobile advertising company, to settle a claim that the company had deceptively tracked the locations of hundreds of millions of consumers in order to display geo-targeted advertisements.[369] The FTC alleged that, despite InMobi's representations that it would only track users' locations after obtaining consent, the company actually tracked location data regardless of whether it had permission to do so, and even when users had specifically opted out.[370] The consent decree required InMobi to pay $950,000 in civil penalties and implement a new privacy program that will be independently audited throughout the next 20 years.[371]

c. Data Breach Guidance

In September 2016, the FTC published the Data Breach Response: A Guide for Business, to advise businesses on how to deal with data breaches.[372] The guide focused on educating businesses on how to secure their operations following a breach. Suggestions include assembling a team of experts (including legal counsel) to conduct a comprehensive breach response, securing physical areas, and stopping additional data loss by taking affected equipment offline immediately. The guide does note, however, that machines should not be turned completely off until forensic experts arrive and presumably have a chance to conduct analyses.

Additionally, the agency offered guidance on how companies can fix vulnerabilities by examining their network segmentation and creating communications plans that reach all affected parties. The parties that should be notified after a data breach include law enforcement, affected businesses and individuals, and even the media. For HIPAA covered entities and their business associates, the Secretary of the U.S. Department of Health and Human Services must also be notified.

Lastly, the guide features a model letter for notifying individuals whose names and Social Security numbers have been stolen. The letter features an optional attachment of a relevant section from
IdentityTheft.gov concerning steps that consumers should take if their Social Security numbers have been exposed. The letter and attachment can be modified depending on the type of personal information that was lost.

**d. Scope of Authority--Common Carriers**

The question of which corporations fall under the FTC's purview was addressed this year by the Ninth Circuit. In a dispute between AT&T and the FTC over AT&T's allegedly deceptive "data throttling," AT&T argued that, as a common carrier, it was not subject to the FTC's authority.[373] Section 5 of the FTC Act exempts common carriers from the FTC's authority to regulate unfair or deceptive business practices.[374]

On August 29, 2016, the Ninth Circuit agreed with AT&T, holding that the company's status as a telecommunications provider placed it under the common carrier exception.[375] The FTC argued that AT&T should not qualify for the exception because the company engaged in non-common carrier activities, such as providing consumers with mobile data or email services.[376] The court rejected this activities-based approach, placing AT&T and other telecom companies beyond the reach of Section 5.[377] In October 2016, the FTC petitioned the Ninth Circuit to rehear the case en banc.[378]

The Ninth Circuit's decision is significant in light of the FCC's Open Internet Order, which reclassified broadband internet service providers as common carriers.[379] Depending on the final outcome of the case, the FTC's jurisdiction could be significantly reduced, shifting some of the cybersecurity regulatory workload to the FCC.

2. **Department of Health and Human Services ("HHS")**

Throughout 2016, HHS was very active in efforts to safeguard patient privacy. On March 21, 2016, HHS began the second phase of its audit program to assess compliance with patient privacy provisions of the Health Insurance Portability and Accountability Act (HIPAA).[380] The audit, which concluded its first phase in 2012, covers both "covered entities," such as health care providers and insurance plans, as well as business associates that handle patient information on behalf of these covered entities.[381] If the audits reveal serious compliance issues, the entity or business associate could be subject to financial penalties and requirements to enter into formal agreements to address the deficiencies.[382]

In 2016, HHS achieved several multimillion dollar settlements over HIPAA violations. The largest settlement occurred in August, when Advocate Health Care System agreed to pay $5.55 million to settle a variety of HIPAA violations.[383] Among the violations was a data breach of Advocate's subcontractor billing company that exposed sensitive patient information.[384] HHS found that Advocate failed to obtain written assurances from its business associate that electronic patient data would be appropriately protected.[385] This settlement, the largest to date against a single entity,[386] as well as other large payouts over HIPAA violations, could lead to increased spending by health care providers on compliance-monitoring and related services.

Additionally, this year, HHS outlined its position on the status of cloud service providers who manage electronic protected health information. In its report titled "Guidance on HIPAA & Cloud Computing,"
HHS confirmed that, outside of very narrow exceptions, these cloud service providers are business associates covered by HIPAA.[387] Thus, these service providers must enter into business associate agreements with covered entities and other business associates prior to handling patient data.

3. Securities and Exchange Commission

a. Making Cybersecurity Examination a Priority

Beginning with the Securities and Exchange Commission's ("SEC") issuance of cybersecurity guidance in 2011, the SEC has continued to increase focus on "assessing the readiness of market participants and providing investors with information on how to better protect their online investment accounts from cyber threats . . . ."[388] In May, at the Reuters Financial Regulation Summit, former SEC Chair Mary Jo White explained that cybersecurity is the biggest risk facing the financial system.[389] White noted that some "major exchanges, dark pools and clearing houses did not have cyber policies in place that matched the sort of risks they faced[.]"[390] Accordingly, White stated that the "SEC can't do enough in this sector[.]"[391] It comes as no surprise then that the Office of Compliance Inspections and Examinations ("OCIE"), a division of the SEC that promotes compliance with securities laws, identified cybersecurity as one of its selected examination priorities for 2016, and once again at the start of 2017.[392] The OCIE noted that as part of the SEC's mission to also maintain "fair, orderly, and efficient markets[,]"[393] the OCIE would be examining structural risks and trends that could potentially "involve multiple firms or entire industries."[394]

In 2015, the OCIE conducted testing focused on risk assessment, access rights and controls, data loss prevention, vendor management, training, and incident response.[395] Building on its efforts in 2015 to examine broker-dealers' and investment advisers' "cybersecurity compliance and controls[,]"[396] the OCIE announced that in 2016 it would "advance these efforts, which include[d] testing and assessments of firms' implementation of procedures and controls."[397] An update on the OCIE’s progress and its findings is forthcoming.

In June, the SEC created a new advisory position, Senior Advisor to the Chair for Cybersecurity Policy.[398] The new position will assist the SEC chair on "all cybersecurity policy matters . . . responsible for coordinating efforts across the agency to address cybersecurity policy, engaging with external stakeholders, and further enhancing the SEC's mechanisms for assessing broad-based market risk."[399] The Senior Advisor will also assist the SEC in enhancing its "coordinated approach to cybersecurity policy . . . and engage at its highest levels with market participants and governmental bodies concerning the latest developments[,]"[400]

b. Enforcement Actions

The SEC has only brought two significant enforcement actions since beginning its cybersecurity examination of broker-dealers and investment advisers in 2014.[401] It resolved the second of these actions on June 8, 2016, when Morgan Stanley agreed to pay a $1 million penalty to settle charges related to its alleged failures to protect customer information, some of which allegedly was hacked and offered for sale online.[402] The SEC pursued this action notwithstanding the fact that Morgan Stanley self-detected the breach during a routine sweep, took prompt corrective actions to remove the stolen data
from the internet, and promptly notified the proper authorities. Further, no investors suffered financial harm. In its order, the SEC found that Morgan Stanley had not "adopt[ed] written policies and procedures reasonably designed to protect customer records and information, in violation of Rule 30(a) of Regulation S-P[",]" the "Safeguards Rule."[403] Specifically, Morgan Stanley allegedly did not conduct any auditing or testing of its "portals" that allowed for access to customer data.[404] Further, Morgan Stanley allegedly did not monitor user activity in these "portals" to help identify unusual or suspicious activity.[405] As a result, a Morgan Stanley employee allegedly was able to impermissibly access confidential customer data and copy the data to his personal server.[406] The employee's server was then hacked, which led to the customer data being posted on the internet.[407]

Despite Morgan Stanley's prompt corrective actions, the SEC still found Morgan Stanley's violation to be willful.[408] Thus, the SEC's punitive measures suggest that it is more interested in whether or not a breach has occurred, not whether any investors have suffered actual financial harm as a result of the breach.[409]

4. Federal Communications Commission ("FCC")

The FCC was especially active in the data privacy and cybersecurity space in 2016. It released a new data privacy regulation, initiated a number of enforcement actions, and provided a declaratory ruling.

a. FCC Rulemaking Regarding ISPs

On October 27, 2016, the FCC voted to adopt sweeping new regulations to govern the ways in which providers of broadband Internet access service ("BIAS") can use and share their customers' proprietary information.[410]

There are three key components to the new rules. First, broadband providers must provide consumers with clear notice of their data collection and use policies. Second, broadband providers must allow consumers to opt out of having "non-sensitive" information used by the providers, or shared by the providers with third parties. Broadband providers must also obtain affirmative opt-in consent before they can use or share "sensitive" customer data, which is defined to include information such as location, health records, and the contents of electronic messages. Third, broadband providers must abide by more stringent and specific requirements for notification of any data breaches.

The most important element is the broad-reaching "consent" requirement. Specifically, providers must obtain express "opt-in" consent before they may use "sensitive" individually-identifiable consumer information, and before sharing that information with third parties.[411] Sensitive information is defined broadly, to include precise geo-location; children's information; health information; financial information; Social Security numbers; web browsing history; mobile application(s) usage history; and the contents of any communications.[412]

In addition, broadband providers must provide consumers with an opportunity to "opt out" of consenting to the use and sharing of their non-sensitive information. Non-sensitive information includes all remaining personally-identifiable information, such as service tier information, that could be used for targeted advertising or other commercial purposes.[413]
The exceptions to the consent requirements are limited. Customer consent is inferred only where non-sensitive information is used and shared for marketing telecommunications-related services, billing and collecting for the broadband provider's services, preventing fraudulent use of the provider's network, and in certain specified emergency situations.[414]

The industry-wide implications of the FCC's new rules are substantial. The rules are likely to cause confusion for broadband providers and other players in the Internet ecosystem given the often-overlapping jurisdiction of the FCC and the FTC, as well as state regulation of related issues including data breach reporting. For example, under the FTC approach, web browsing history and application usage data is not considered sensitive information. Commissioner Michael O'Reilly, in his dissenting statement to the new rules, argued that "[r]equiring opt-in consent for these categories will…upend years of settled expectations, burdening rather than benefitting most users."[415]

Industry reaction to the new rules has mostly been negative. A spokesman for the Direct Marketing Association, a trade group dedicated to data-driven marketing, commented that requiring a consumer's opt-in consent before certain data can be used or shared will "unnecessarily disrupt the advertising ecosystem that fuels the explosive growth of the online economy."[416] On the other hand, The Center for Democracy and Technology, a consumer privacy group, commended the rules as a "significant step forward in protecting internet users, who have no choice but to expose massive amounts of information to broadband providers."[417]

In addition, the new rules do not apply to "edge providers," which the FCC defines as any "individual or entity that provides any content, application, or service over the Internet," or that provides a device used to access content, applications, or services over the Internet.[418] As a result, edge providers--including, for example, many social media services--are not constrained under these regulations in their collection or use of consumer information. Moreover, the rules present many practical challenges that broadband providers should consult closely with counsel to address, including whether it is necessary or sufficient to obtain opt-in or opt-out consent for each instance of data use, or whether a blanket consent (included, for example, in a user agreement) will suffice.

b. Data Security Settlements

In 2015 and 2016, the FCC entered into consent orders with a number of companies for purportedly violating the Communications Act of 1934, mostly by failing to properly protect customers' personal information.

In 2015, the FCC reached three major settlements with telecommunications companies for purportedly failing to adequately protect and secure their customers' personal information (including names, Social Security numbers, and account-related data) under Section 222 of the Communications Act. The total value of these settlements equaled just under $30 million.[419]

Most recently, on March 7, 2016, the FCC reached a $1.3 million settlement with Verizon Wireless to resolve its investigation into whether Verizon broke data security rules with its "supercookies" advertising program.[420] Supercookies create unique identifier headers ("UIDH"), which are pieces of software that track customers' Web usage, used to identify customers in order to deliver targeted
In addition to the civil fine, Verizon has agreed to notify customers of its advertising program and to obtain opt-in consent before sharing this type of information with third parties and will obtain customers' opt-in or opt-out consent before sharing UIDH internally within the Verizon corporate family.[422]

5. Consumer Financial Protection Bureau

Like the FCC, the Consumer Financial Protection Bureau ("CFPB") flexed its regulatory muscles in 2016 by engaging in both rulemaking and enforcement related to privacy and cybersecurity.

a. CFPB Amendment to Annual Privacy Notices

In July 2016, the CFPB proposed a new amendment to Regulation P of the Gramm-Leach-Bliley Act ("GLBA"). This rulemaking would implement a December 2015 statutory amendment to the GLBA exempting certain financial institutions from the requirement to provide annual privacy notices.

By way of background, the GLBA and Regulation P "mandate that financial institutions provide their customers with initial and annual notices regarding their privacy policies."[423] These required privacy notices are intended to provide customers with information regarding how a financial institution shares their nonpublic personal information, including personally identifiable financial information, with other entities.[424] In some cases, these notices also explain how consumers can opt out of certain types of sharing.[425]

Currently, Regulation P contains an exception whereby a financial institution has the option to post its annual privacy notice on its website (known as the "alternative online delivery method"), if it meets a number of requirements which relate to certain types of information-sharing activities.[426] Otherwise, it must send the privacy notice through U.S. postal mail.

This amendment would take the alternative online delivery method one step further, by exempting certain financial institutions--who, as with the alternative online delivery method, meet particular conditions--from the requirement to provide annual privacy notices altogether.[427] A financial institution is not required to provide an annual notice if it: (1) provides nonpublic personal information only in accordance with certain exceptions in the GLBA concerning consumer opt-out rights and (2) has not changed its policies and practices with regard to disclosing nonpublic personal information from the policies and practices that were disclosed in its most recent notice sent to consumers.[428] This exception would replace the alternative online delivery method.

b. Dwolla Inc. Enforcement Action

The CFPB engaged in its first-ever enforcement action related to data security in March 2016, when it entered a consent order against online payment platform Dwolla, Inc. ("Dwolla").[429] In short, the CFPB found that Dwolla engaged in deceptive acts by making materially false statements concerning its data security practices and the safety of its online payment system. The CFPB noted that Dwolla made a number of representations to its customers relating to its data security practices, including that its
practices "exceed industry standards" and that all customer information is "encrypted and stored securely." The CFPB found that Dwolla's practices, in reality, fell short of these claims.

While the civil penalty was relatively small ($100,000), the CFPB imposed fairly intrusive remedial measures that require Dwolla to undertake a long list of actions to improve its data security practices. Among other things, Dwolla must develop new data security practices and procedures; conduct regular, mandatory employee training on data security; and retain an independent person to conduct a data security audit. Notably, this enforcement action was pursued despite the fact that the CFPB did not allege that Dwolla's systems had been breached.

The CFPB brought this enforcement action under its Dodd-Frank authority to regulate parties who engage in unfair, deceptive, and abusive acts and practices ("UDAAP") in connection with consumer financial products and services. Relative to earlier interpretations of the scope of the CFPB's authority, this is a clear expansion of the UDAAP provision and the CFPB's jurisdiction.

It remains to be seen whether the Dwolla enforcement action represents a significant shift in the CFPB's regulatory focus. While it could be an isolated matter, it could also signify the CFPB's arrival as a player in data privacy regulation and enforcement going forward.

6. State Attorneys General

During the past year, state attorneys general offices continued to develop privacy and cybersecurity regulation, bring enforcement actions, and issue guidance. While California and New York are among the most aggressive states, and therefore remain at the forefront of developments in this area, other states are also taking steps to ensure compliance with state and federal laws related to data privacy and cybersecurity. We will likely see more activity from state attorneys general in the coming years, especially if the Trump administration does not make data privacy and cybersecurity a top concern at the federal level.

a. California

In 2016, the California Attorney General's Office ("California AG") continued its practice of issuing substantive reports and guidance documents for both consumers and businesses. Among the most significant is the Data Breach Report released in February 2016, which contained, for the first time, a list of twenty "critical security controls" developed by the Center for Internet Security that constitute the "minimum level of information security" acceptable for entities that collect or handle PII. Among these controls are "Inventory of Authorized and Unauthorized Devices," "Secure Configurations for Hardware and Software," and "Continuous Vulnerability Assessment and Remediation." The Data Breach Report takes the position that noncompliance with these twenty minimum standards amounts to a "lack of reasonable security" under California Civil Code § 1798.81.5 and possibly—it remains to be seen—under common law negligence as well. This could influence future interpretations of California's data privacy legislation and may serve as a model for other law-making and regulatory bodies. In addition, the Data Breach Report contains an assessment of the types of data most likely to be breached. For example, the report recommends that organizations limit their collection of Social Security numbers...
because (i) this information is among the categories of data most likely to be breached and (ii) breaches of social security numbers are among the most damaging to consumers.[434]

In addition to the Data Breach Report, the California AG recently issued Ready for School, a guidance document addressed to the education technology industry and calling for stronger protection of K-12 students' personal data.[435] Under both state and federal law, entities that handle student data must meet particularly rigorous privacy standards, and the California AG's guidance document is intended to help schools and businesses understand this particular legal context.[436]

The California AG has also taken steps to make reporting of alleged privacy violations easier for consumers, including the launch of a streamlined online form to report violations of the California Online Privacy Protection Act ("CalOPPA").[437] For example, a consumer could use this form to report that a website has failed to post adequate privacy policies or has ignored a "do not track" request.[438]

The California AG was less active during the past year on the litigation front, but did announce one settlement in connection with allegedly unlawful call recording.[439]

b. New York

In comparison to California, the New York Attorney General's Office ("New York AG") in 2016 was more active in litigation and less active in the development of industry- and consumer-oriented guidance. The New York AG Office, led by Eric Schneiderman, settled several cases during the past year alleging violations of state privacy and data breach notification laws. For example, the New York AG settled a case against then presidential-nominee Donald Trump's hotel chain arising from a series of malware-enabled breaches that occurred in 2014 and 2015, which the chain allegedly failed to report for several months in violation of New York law.[440] The New York AG also settled a case against EzcontactsUSA, alleging that the online contact-lens retailer misrepresented the security of its website, failed to secure customers' payment information, and neglected to report a data breach once discovered.[441] While the settlement amounts in the Trump hotels and EzcontactsUSA cases are relatively modest, only $50,000 and $100,000 respectively, both settlements require the defendants to take specific steps to strengthen their data security practices.[442]

The New York AG also took action in response to alleged violations of federal law, most notably COPPA, which prohibits the collection of user data on websites intended for users under the age of 13. In September, the New York AG announced the conclusion of an investigation it named "Operation Child Tracker," which probed the online privacy and data collection practices of Viacom, Hasbro, Mattel and JumpStart Games and advertisers whose content they hosted on their websites.[443] In settling the investigation, the companies agreed to pay a combined penalty of $800,000 and to enact "comprehensive reforms" of their advertising practices.[444] Hasbro, which participated in a FTC-approved safe harbor program, did not pay a penalty.[445]

Last January, the New York AG also resolved litigation with ride-sharing service Uber. The New York AG alleged that Uber displayed riders' personal information in an aerial view (known internally as the 'God View') and left the data vulnerable to third parties. While the settlement amount was again quite
small, Uber agreed to encrypt riders' GPS information and limit its employees' access to it and other sensitive data.[446]

c. Developments in Other States

The Attorney General for the state of Washington published the state's first data breach report in September 2016, approximately one year after its updated data breach notification law was signed into law.[447] In Texas, Attorney General Ken Paxton reached a settlement with an app developer based in California, Juxta Labs, in a case alleging that Juxta's child-friendly apps transmitted users' GPS coordinates and IP addresses in violation of COPPA.[448]

States have also continued to coordinate their enforcement strategies with each other. Last November, the attorneys general of 15 states teamed up to resolve a joint investigation of Adobe, Inc., initiated after the software developer's 2013 data breach. Abode agreed to pay a total of $1 million and to take steps to prevent future breaches.[449] And in December, 13 states and the District of Columbia joined forces with the Federal Trade Commission ("FTC") to command a $1.6 million settlement from the operator of AshleyMadison.com, which was hacked in 2015, resulting in the exposure of 36 million users' (particularly) personal information.[450] The FTC's lawsuit alleged that AshleyMadison had not only failed to protect the data it collected, it also misled consumers about the safety of their data. The settlement requires the website operator to implement data-privacy controls and undertake regular security audits, and suspends an additional $8.75 million judgment due to AshleyMadison's inability to pay.

7. New York Department of Financial Services ("NYDFS")

On September 13, 2016, New York staked its position at the front lines of cybersecurity protection, when the NYDFS proposed first-in-the-nation cybersecurity rules for banks, insurers, and financial services companies. Announcing the proposal, New York Governor Cuomo stated, "This regulation helps guarantee the financial services industry upholds its obligation to protect consumers and ensure that its systems are sufficiently constructed to prevent cyberattacks to the fullest extent possible."[451] The proposed rules, now scheduled to go into effect on March 1, 2017, two months later than originally scheduled, will require regulated financial institutions to meet a set of specific standards as set forth by the NYDFS.

Since 2013, the NYDFS has shown an increased interest in cybersecurity, gathering information on cybersecurity practices and incidents in the banking sector. From May 2014 through April 2015, the NYDFS issued a series of cybersecurity reports in the banking and insurance sectors.[452] During that time, the NYDFS has also been collecting information from financial and insurance institutions, meeting with cybersecurity experts, and conducting surveys to inform its latest proposal.

As proposed, 23 NYCRR Part 500 would require regulated financial institutions to establish cybersecurity programs, adopt written cybersecurity policies, appoint an internal Chief Information Security Officer ("CISO"), and have policies and procedures designed to safeguard information accessible to third parties, along with a variety of other requirements to protect the confidentiality and integrity of information systems. Such requirements include, but are not limited to:
• Annual penetration testing and vulnerability assessments;

• Limitations and periodic reviews of access privileges;

• Written application security procedures, guidelines and standards that are reviewed and updated by the CISO at least annually;

• Annual risk assessment of the confidentiality, integrity, and availability of information systems; adequacy of controls; and how identified risks will be mitigated or accepted;

• Employment and training of cybersecurity personnel to stay abreast of changing threats and countermeasures;

• Monitoring of authorized users and cybersecurity awareness training for personnel;

• Encryption of all nonpublic information held or transmitted; and

• Written incident response plan to respond to, and recover from, any cybersecurity event.[453]

The proposed rule stems from a growing concern to protect financial institutions from cyberattacks, such as those that have been recently perpetrated against some of the world's biggest banks, such as JP Morgan Chase and Wells Fargo. While other regulatory bodies have issued similar guidance, New York was the first to propose mandatory regulations in this sphere. About one month after the NYDFS' proposal, on October 19, 2016, federal banking regulators, including the Federal Reserve, Office of the Comptroller of the Currency, and the Federal Deposit Insurance Corporation, released an advance notice of proposed rulemaking that would impose heightened cybersecurity standards on many of the same financial institutions.[454]

The proposed regulations were revised after DFS received over 150 comments, many of which criticized their broad application to companies of all sizes. Rejecting this critique, the regulations still will apply to any companies "required to operate under a license, registration, charter, certificate, permit, accreditation or similar authorization under the banking law, the insurance law or the financial services law."[455] The most notable change in the revised regulations is the incorporation of an expanded transitional period for regulated entities to become compliant with the regulations.[456]

Other modifications relax some of the original proposed regulations, including, in some cases, linking them to each regulated entity's own periodic Risk Assessment. Regulated entities are now to perform a Risk Assessment based on the individual characteristics of the companies to determine how to formulate its cybersecurity program to comply with the regulations. Each entity's own Risk Assessment will then help inform how it can comply with the proposed regulations. For example, instead of universally mandating penetration testing and vulnerability assessments on a quarterly basis, the revised rules require covered entities to include monitoring and testing developed in accordance with its Risk Assessment. Risk Assessments will now be performed only "periodically," rather than annually as originally proposed.[457]
GIBSON DUNN

Notably, DFS added a materiality requirement to the breach-reporting requirement in the proposed regulations. The mandatory reporting provision requiring regulated entities to notify DFS within 72 hours of any Cybersecurity Event that involved the "actual or potential unauthorized tampering with, or access to or use of, Nonpublic information" has been revised to require notification within 72 hours only for "Cybersecurity Events that have a reasonable likelihood of materially harming any material part of the normal operation(s) of the Covered Entity," or are otherwise required to be reported to a governmental body.[458] Other changes include a narrowed definition of "Nonpublic information."[459]

While the changes reflect an overall attempt to provide more time and flexibility for regulated entities to come into compliance, the regulations still impose considerable requirements and burdens on a wide range of companies.

8. Other Agencies

a. Cybersecurity National Action Plan and the Commission on Enhancing National Cybersecurity

On February 9, 2016, President Obama's administration unveiled the Cybersecurity National Action Plan ("CNAP").[460] The CNAP is designed to "put[] in place a long-term strategy to enhance cybersecurity awareness and protections, protect privacy, maintain public safety as well as economic and national security, and empower Americans to take better control of their digital security."[461] As part of these goals, the CNAP includes the establishment of the Commission on Enhancing National Cybersecurity ("CENC"),[462] as well as the proposal of a $3.1 billion Information Technology and Modernization Fund that fosters the creation of a new Federal Chief Information Security Officer position.[463] Further, the administration will invest over $19 billion for cybersecurity as part of the President's Fiscal Year 2017 budget, representing a more than 35 percent increase from Fiscal Year 2016.[464]

In furtherance of CENC's goals to make detailed recommendations to "strengthen cybersecurity in both the public and private sectors while protecting privacy,"[465] on April 13, 2016, the President selected 12 individuals to serve on the commission.[466] The individuals include company executives, professors, research scholars, and former national advisors, among others.[467] President Obama made it clear that cybersecurity is one of the most important challenges the nation faces, and that the 12 individuals on the commission are charged with "recommending bold, actionable steps that the government, private sector, and the nation as a whole [could] take to bolster cybersecurity in today's digital world[.]")[468]

On December 1, 2016, CENC presented its findings to President Obama.[469] In its report, it noted six priority imperatives and 16 recommendations.[470] The imperatives include, among others, preparing consumers to thrive in a digital age, building cybersecurity workforce capabilities, and ensuring a fair, competitive, and secure "global digital economy."[471] Within each recommendation, the commission also suggested action items for the administration. Of note, it recommended that the incoming administration create an Assistant to the President for Cybersecurity, who would report to the National
b. Department of Energy Cybersecurity Infrastructure Funding

The Department of Energy's Office of Electricity Delivery and Energy Reliability ("OE") has recognized that enhancing cybersecurity is a crucial aspect of protecting the nation's power grid.[475] Because all sectors of the nation depend on the electrical grid infrastructure, protecting this infrastructure from cyberattacks is particularly important.[476]

To further its cybersecurity goals, on January 20, 2016, the OE announced approximately $23 million in funding for research and development of advanced cybersecurity technologies.[477] The funding is designed to address challenges in power grid modernization, and to maintain sufficient scientific advances in order for the energy sector to continue adapting to the ever changing cyber landscape.[478] Following this announcement, in July, the Department of Energy announced additional funding to further support and protect the electric grid from attacks.[479] As part of this additional funding, the Department will provide up to $15 million, subject to congressional appropriations, to support efforts by the American Public Power Association and the National Rural Electric Cooperative Association to "further enhance the culture of security within their utility members' organizations."[480] This funding will also be used to "develop security tools, educational resources, updated guidelines, and training on common strategies."[481]

In August, the Department awarded up to $34 million in funding through the Cybersecurity of Energy Delivery Systems Program to 12 projects that represent energy sector organizations.[482] The 12 projects are designed to "enhance the reliability and resilience of the nation's energy critical infrastructure through innovative, scalable, and cost-effective research, development and demonstration of cybersecurity solutions."[483]

c. Federal Deposit Insurance Corporation Rulemaking on Enhanced Cyber Risk Management Standards

In October, the Federal Deposit Insurance Corporation, along with the Federal Reserve Board and the Office of the Comptroller of the Currency, approved an advance notice of proposed rulemaking to invite comments on potential heightened cybersecurity standards.[484] The heightened standards will apply to the largest and most interconnected entities such as financial institutions with total assets of $50 billion or more, and non-U.S. banks with total U.S. assets of $50 billion or more.[485] Further, the agencies will consider applying the heightened standards to certain third-party services as well.[486] The standards imposed on applicable entities are "aimed at increasing [ ] operational resilience and reducing
B. Legislative Developments

1. Federal Developments

In 2016, Congress introduced a number of pieces of proposed legislation related to cybersecurity and data privacy. The proposed legislation spans a wide range of topics, but principally concerns three key categories: international data privacy, cybersecurity preparedness, and cybersecurity disclosures. Despite this heavy influx of proposed cyber legislation, in 2016, President Obama only signed one such bill into law.

a. International Data Privacy Law

The single cyber-related bill that made it to the President's desk in 2016 is the Judicial Redress Act of 2015. The law extends the protections of the Privacy Act of 1974 to the citizens of "covered" foreign countries. Accordingly, the law gives certain foreign citizens (primarily countries in the EU) the right to seek redress in U.S. courts for privacy violations when their personal information is shared with law enforcement agencies. Pursuant to this law, the Attorney General has authority to add countries to the "Privacy Act List"; however, the law qualifies this authority by requiring that, prior to adding a country to the list, the Attorney General certify that: (1) the country entered into an agreement with the U.S. that provides privacy protections for information shared for the purpose of preventing, investigating, detecting, or prosecuting criminal offenses, or has "effectively shared" such information with the U.S.; (2) the country permits the transfer of data for commercial purposes between its territory and the U.S.; and (3) this data-transfer agreement does not "materially impede the national security interests of the U.S." President Obama signed the bill into law on February 24, 2016. The U.S. tech industry was a major backer of the legislation, describing it as a "critical step in rebuilding the trust of citizens worldwide in both the U.S. government and [the tech] industry." The tech industry's recent focus on instilling "trust" arises in part from its concern for the broad international ramifications of Edward Snowden's highly publicized disclosures regarding the role of many leading tech companies in NSA surveillance activities.

b. Cybersecurity Preparedness

Congress introduced a number of bills relating to cybersecurity preparedness. One example is the National Cybersecurity Preparedness Consortium Act of 2016, which the House passed on May 16, and is now in the Senate's Committee on Homeland Security and Governmental Affairs. Representative Joaquin Castro (D-TX), who introduced the bill, said that its purpose is to "allow[] the Department of Homeland Security ("DHS") to collaborate with experts outside of the government to improve state and local cyber preparedness." To accomplish this, the bill would authorize the DHS to work alongside a consortium, composed of primarily nonprofit entities, to support efforts to address cybersecurity risks and incidents. The DHS, together with the consortium, would be permitted to conduct cross-sector cybersecurity training and simulation exercises, aid states in developing cybersecurity information...
sharing programs, and assist in incorporating cybersecurity risk prevention and response into existing state emergency plans, among other things.[497]

A second bill implicating the DHS's role in cybersecurity preparedness is the Cyber Preparedness Act of 2016. The purpose of the bill is to "enhance preparedness and response capabilities for cyber attacks and bolster the dissemination of homeland security information related to cyber threats."[498] This bill would amend the Homeland Security Act of 2002 to require DHS's State, Local, and Regional Fusion Center Initiative to coordinate with the National Cybersecurity and Communications Integration Center ("NCCIC") to provide fusion centers with DHS cybersecurity resources.[499] The bill would require the DHS to support fusion centers, review cybersecurity risks gathered by fusion centers, and disseminate cybersecurity risk information to fusion centers.[500] The House passed the bill on September 26 and the Senate referred it to the Committee on Homeland Security and Government Affairs on September 27, 2016.[501]

The Improving Small Business Cybersecurity Act of 2016 also focused on preparedness. The bill would amend the Small Business Act to allow the Small Business Administration ("SBA") to make grants to small business development centers ("SBDCs") so they can assist small businesses in improving preparedness against cyber threats.[502] To achieve this preparedness goal, the SBDCs would distribute cybersecurity risk information to help small businesses develop cybersecurity infrastructure, threat awareness, and employee training programs.[503] To qualify as a "small business" under the Small Business Act, there are two widely used size standards: (1) a maximum of 500 employees for most manufacturing and mining industries, or (2) a maximum of $7.5 million in average annual receipts for many nonmanufacturing industries.[504] While this is the general rule, there are several industry-specific exceptions.[505] Commenting on the importance of the bill, Representative Steve Chabot (R-OH), House Small Business Committee Chairman, noted that "American small businesses are under cyberattack like never before," and that Congress needs to "do[] all [they] can to help protect [] job creators and their customers against the great and growing array of cyber-threats they face on a daily basis."[506] The House passed the bill on September 21, 2016, and the Senate received it on September 22, 2016.[507]

c. Cybersecurity Disclosures

Responding to the increasing prevalence of data breaches, the Cybersecurity Disclosure Act of 2015 "asks publicly traded companies to include information pertaining to cybersecurity in their Security Exchange Commission ("SEC") filings."[508] The goal of the bill is to "promote transparency in the oversight of cybersecurity risks at publicly traded companies."[509] The bill would require publicly traded companies to disclose, in their annual report or proxy statement, whether any member of the company's governing body is a "cybersecurity expert," or whether any member of the governing body has "experience" in cybersecurity.[510] If there is no board member with cybersecurity expertise or experience, the bill would require the company to explain, in its disclosures, what additional measures it took regarding cybersecurity.[511] Harvard Law School Professor John Coates, a supporter of the bipartisan bill, argues that "[it] would encourage boards to take direct responsibility for cybersecurity through a light touch 'comply or disclose' approach, preserving flexibility for companies to respond to
cyber threats in a tailored and cost-effective way."[512] The Senate referred the bill to the Committee on Banking, Housing, and Urban Affairs on December 17, 2015.[513]

d.  Data Breach Notification

Despite progress in other areas, Congress has not passed a federal data breach notification law. Accordingly, the notification requirements associated with data breaches vary among the 47 states that have adopted laws on the subject.[514] There was some movement toward federalizing data breach notification requirements in 2015, with the introduction of the Personal Data Notification and Protection Act of 2015; however, the bill remains pending in the Subcommittee on the Constitution and Civil Justice.[515] The bill would require certain businesses that use, access, transmit, or store "sensitive personally identifiable information about more than 10,000 individuals during any 12-month period" to notify individuals whose information is believed to have been accessed or acquired through a discovered security breach.[516]

e.  Cybersecurity under Trump

During a campaign speech in October, then presidential-nominee Donald J. Trump described cybersecurity as a "top priority" for his future administration.[517] Now that President Trump is a reality, many are wondering what actions he will take on matters of cybersecurity. One possibility is that Trump will support the passage of an encryption bill. In April 2016, Senators Diane Feinstein (D-CA) and Richard Burr (R-NC) introduced the draft for a bill called "The Compliance with Court Orders Act of 2016," which would require providers of communication services and software to give "responsive, intelligible information or data, or appropriate technical assistance" to the government pursuant to a court order.[518] When introduced, President Obama refused to support the legislation.[519] It remains to be seen how President Trump will react to such a bill in 2017.

2.  State Developments

In 2016, at least 26 states passed legislation related to cybersecurity. Common themes among these various state laws include: increased reporting requirements for cybersecurity incidents and expenses, establishing committees on cybersecurity, amending public records laws, criminalizing certain cyber conduct, and imposing data breach notification requirements.

a.  Reporting Requirements

At least two states passed laws that involve the reporting of certain cyber-related activities. California, for example, enacted a law that requires state agencies to give the Department of Technology a report on actual and projected cybersecurity expenses.[520] Oregon also passed a law related to reporting requirements, which mandates that state agencies promptly notify the Legislative Fiscal Office with information following any data security incidents.[521] The Oregon law further imposes a requirement that state agencies compile annual reports on state information security.[522]
b. State Committees

The establishment of state committees on cybersecurity was another theme in 2016, with at least four states adopting laws on the subject. Colorado established the "Colorado Cybersecurity Council," which aims to develop cybersecurity policy and providing guidance to the governor on cyber-related issues. Similarly, the legislature in Georgia passed a law creating the "Senate Data Security and Privacy Study Committee," which aims to evaluate Georgia's current data security procedures and identify any existing or potential vulnerabilities.

c. Public Records

Another trend in 2016 saw state legislatures limiting the definition of "public records" for security purposes. The Delaware legislature passed a law that amends the State Freedom of Information Act to exclude information on technical infrastructure, and related details, from the public record. The purpose of this exclusion is to ensure the security of the state's information and technology system. Similarly, the Florida legislature adopted a law that exempts information that is related to the state's technology systems from the public record. Virginia also passed a law that excludes information related to cybersecurity from the public record.

d. Criminalization of Cyber Conduct

A number of states passed laws defining the criminality of certain cyber conduct. In California, a new law makes it a crime for a person to knowingly infect a computer, computer system, or computer network with ransomware. Steve Giles, the CIO of Hollywood Presbyterian Medical Center ("HPMC"), offered testimony before the state senate's Public Safety Committee in which he recounted the night of February 5, 2016, when a ransomware attack made "[e]very system within the medical center [ ] inaccessible." Shortly thereafter, HPMC received and paid ransom demands amounting to $17,000 in order to recover its files from the attack's perpetrator. Following the HPMC incident, numerous other California hospitals fell victim to ransomware attacks. Commenting on the need to criminalize this conduct, Senator Bob Hertzberg (D-Los Angeles), who introduced the bill, noted that it is critical to "have an up-to-date law that works practically in the system of justice to deal with this new ransomware threat."

Taking a more comprehensive approach to criminalizing cybercrime, Washington adopted the State Cyber Crime Act, which criminalizes "computer trespass, electronic data service interference, spoofing, electronic data tampering, and electronic data theft." Representative Chad Magendanz (R-Issaquah), who introduced the bill, suggests that its importance transcends the State of Washington, noting that "[o]ther states will be using [the State Cybercrime Act] as a model." Magendanz came to Washington politics from a career at Microsoft, where he was a "key player" in identifying and preventing cybersecurity threats. His experience at Microsoft led Magendanz to propose a piece of legislation that "effectively targets true criminals." Magendanz argues that, by expressly criminalizing behavior, rather than attempting to regulate the technology itself, the new law will "eliminate barriers for law enforcement" and "allow prosecutors to go after [ ] criminals before they've stolen [ ] data."
e. Data Breach Notification

While there is no uniform federal data breach notification law, 47 states, as well as the District of Columbia, Guam, Puerto Rico, and the Virgin Islands have enacted legislation requiring some form of notification following a breach.[539] With Tennessee adopting their notification law in 2016, Alabama, New Mexico, and South Dakota are now the only states without laws on the subject.[540] The Tennessee law adds complexity to the already conflicting requirements of the other 46 states. Unlike most other state laws, Tennessee's legislation requires notification not only for the loss of unencrypted data, but also for encrypted data that includes personally identifiable information.[541]

III. U.S. Government Data Collection

A. Microsoft's Challenge to "Gag Orders"

In April 2016, Microsoft took the technology sector's effort to limit government access to user data on the offensive, challenging the government's authority to apply for nondisclosure or "gag" orders, which prevent cloud storage companies from disclosing government seizures of user data.[542] After receiving thousands of such orders over the past two years, Microsoft sued the Department of Justice in the Western District of Washington, alleging that 18 U.S.C. §§ 2703 and 2705(b)--which permit these orders--violate the Constitution; specifically, that section 2703 violates the Fourth Amendment, and section 2705(b) violates the First and Fourth Amendments.[543] Under section 2703(b)(1)(A), the government can obtain a warrant to seize electronic communications held by a cloud storage provider, such as Microsoft, without notifying the user whose data is seized. Section 2705(b) further allows the government to apply for a nondisclosure order barring the storage provider itself from disclosing the existence of the warrant. A court must issue such an order if it finds "reason to believe" that disclosure of the search warrant will endanger public safety, jeopardize an ongoing investigation, or unduly delay trial.[544] A section 2705(b) nondisclosure order may last "for such period as the court deems appropriate."[545] In its first amended complaint, Microsoft alleged that of the over 3,250 nondisclosure orders that it has received since May 2016, nearly two-thirds of them had no fixed end date, meaning that many users may never learn that their data has been searched.[546]

Microsoft claimed that by preventing, ex ante, "speech about the government's access to customers' sensitive communications and documents," Section 2705(b) nondisclosure orders operate as both prior restraints and content-based restrictions on its speech.[547] They violate the First Amendment, Microsoft argued, because they are not narrowly tailored and because the government does not have to show a specific compelling interest in applying for them.[548] Microsoft also alleged that Section 2705(b) is unconstitutionally overbroad, as the "reason to believe" standard does not require narrow tailoring and may be met by a vague showing that disclosure would otherwise seriously jeopardize[e] an investigation or unduly delay[] a trial," and because nondisclosure orders may be of indefinite duration.[549] Finally, Microsoft brought a Fourth Amendment claim on behalf of its customers whose data has been seized. It contended that for a search to be reasonable under the Fourth Amendment, the government must give notice to the target of the search, and that sections 2703 and 2705(b) thus violate the Fourth Amendment by permitting that notice to be delayed indefinitely.[550]
The government filed a motion to dismiss attacking each of Microsoft's substantive claims and its standing to bring them. First, the government argued, Microsoft's facial overbreadth challenge fails because it has not alleged that any particular application of Section 2705(b) is unconstitutional. Furthermore, the "reason to believe" standard, the government's compelling interest in the integrity of investigations, and the use of indefinite nondisclosure orders are all supported by case law. Second, the government contended that Microsoft has a reduced First Amendment interest in information obtained solely through a government investigation, and that Section 2705(b) nondisclosure orders are thus not "typical prior restraints" subject to heightened scrutiny. Even if they were prior restraints, they are permissible because no nondisclosure order may issue without prior judicial review. The government did not dispute that Section 2705(b) orders are content-based speech restrictions, but argued that they are justified because they serve a well-recognized interest in protecting public safety and the integrity of criminal investigations, and prohibit only speech about warrants issued under Section 2703. Finally, the government contended that third parties, such as cloud storage users, are not entitled to legal notice under the Fourth Amendment, arguing that no such requirement is established in the law, and that Section 2705(b) already provides advance review by a neutral magistrate, the "highest protection available under the Fourth Amendment."

The government also challenged Microsoft's standing to bring both its own First Amendment claim and its claim based on its customers' Fourth Amendment rights. According to the government, Microsoft has not alleged that any individual nondisclosure order is unconstitutional. Because the appropriateness of each such order must be considered in light of case-specific facts, Microsoft failed to identify a sufficiently concrete injury to support standing. Microsoft responded that it is harmed by every Section 2705(b) nondisclosure order it receives, and has thus alleged "thousands of concrete, particularized injuries, both actual and imminent," sufficient to justify standing. The government also challenged Microsoft's standing to assert its customers' Fourth Amendment rights. But Microsoft argued that it has third-party standing because nondisclosure orders undermine user trust, and because nondisclosure orders prevent users from asserting their own rights by preventing them from knowing when their data has been seized. The government also argued that Fourth Amendment rights cannot be vicariously asserted, and that even if they could, Microsoft's relationship to its users is insufficiently close, and its claimed "customer trust" injury too "ethereal," to confer third-party standing.

Numerous amici have supported Microsoft, filing eight amicus briefs representing more than 60 groups. The largest of these is comprised of media organizations, including National Public Radio and The Washington Post, who argue that nondisclosure orders violate their First Amendment rights to receive information, thereby impeding their vital societal role of shedding light on government operation. A group of technology companies including Amazon and Google separately argued that the "gag order" tool is a "troubling outlier" to standard notice requirements, and that the court should follow Supreme Court precedent by recognizing that Fourth Amendment jurisprudence must evolve in step with technology. Microsoft has even received support from a collection of former law enforcement officials--including former United States Attorneys from the Western District of Washington--who argue that notice to the targets of search warrants is vital to public trust in law enforcement and does not undermine law enforcement officers' ability to do their jobs. A group of law professors argued that indefinite gag orders violate both the historical underpinnings and modern interpretations of the Fourth Amendment by denying notice to parties whose information has been
seized.[564] In addition, a group of business organizations and major businesses including The Chamber of Commerce of the United States and The National Association of Manufacturers argued that allowing surreptitious searches of electronically stored data without notice will make users hesitant to store their data in the cloud, thereby harming business by impeding the adoption of valuable cloud storage technology.[565]

Oral argument on the government's Motion to Dismiss was held on January 23, 2017. Although the parties' Joint Status Report and Discovery Plan proposed a stay on discovery pending a decision on that motion, District Court Judge James Robart nonetheless set August 14, 2017 as the deadline for discovery to be completed.[566] All dispositive motions must be filed by September 12, 2017, and trial is set for December 11, 2017.[567]

B. Microsoft's Challenge to Warrant for Emails in Ireland

In 2013, the United States District Court for the Southern District of New York issued a warrant under Section 2703 of the SCA,[568] compelling Microsoft to produce the contents of a customer's email account. Microsoft turned over account information stored in the United States, but refused to turn over the actual emails, which were being stored in Ireland, and moved to quash the warrant to the extent it directed Microsoft to produce such content located abroad. Microsoft argued that federal courts do not have authority to issue warrants for the search and seizure of property located outside the United States,[569] and that the government should request the emails from Irish authorities under the Mutual Legal Assistance Treaty adopted by the two countries in 2001 (the "MLAT"). The Department of Justice argued that it was not required to follow the MLAT process because Microsoft is based in the United States, and that, even though the SCA uses the term "warrant" and the document at issue was labeled a "warrant," it requires, similar to a subpoena, "the recipient to produce information in its possession, custody, or control regardless of the location of that information."[570] On April 25, 2014, the district court held in favor of the government, denying Microsoft's motion to quash, finding Microsoft in contempt, and compelling Microsoft to comply with the warrant.

The Second Circuit reversed on appeal earlier this year.[571] The Court first emphasized the presumption against extraterritoriality that applies when interpreting U.S. laws, which rests on the perception that "Congress ordinarily legislates with respect to domestic, not foreign matters."[572] With that presumption in mind, the court turned to the SCA itself, noting that the SCA (i) contains no explicit or implicit reference to extraterritorial application, and (ii) uses the term "warrant," a term of art with traditional, domestic connotations.[573] The Court also looked to the SCA's legislative history, which confirms that protecting user privacy was Congress's focus, and that, in regard to governmental access, Congress was seeking "to ensure that the protections traditionally afforded by the Fourth Amendment extended to the electronic forum."[574] The purpose of the Fourth Amendment, according to the Court, was to "restrict searches and seizures which might be conducted by the United States on domestic matters."[575] Thus, the Second Circuit held that the SCA does not authorize a U.S. court to issue and enforce an SCA warrant against a U.S.-based service provider for the contents of a customer's electronic communications stored on servers abroad, and that the SCA warrant to Microsoft therefore had no power to compel production of a customer's emails stored in Ireland.[576] The Second Circuit vacated the district court's finding of contempt against Microsoft and remanded the case with instructions for the
district court to "quash the warrant insofar as it demands user content stored outside of the United States."[577]

This was a significant win for Microsoft and the almost 100 companies that filed amicus briefs in support of Microsoft, warning that if forced to comply with the warrant, it would lead to a "global free-for-all" and an evisceration of personal privacy.[578]

C. Amendments to Rule 41 of the Federal Rules of Criminal Procedure

The Second Circuit was not alone this year in considering the government's power to search and seize electronic data across jurisdictional lines. As recommended by the Judicial Conference of the United States, the Supreme Court submitted to Congress on April 28, 2016 proposed amendments to Rule 41 of the Federal Rules of Criminal Procedure.[579] The amendments, which took effect on December 1, 2016, address two issues: (1) access to a device at an unknown location, and (2) access to multiple computers in multiple districts.[580]

First, where a suspect has masked the true location of his or her computer through the use of technology such as Tor, botnets, malware, or other anonymizing software, new Rule 41(b)(6)(A) empowers a judge "with authority in any district where activities related to [the] crime may have occurred" to issue a warrant to use remote access within or outside that district. This is a notable expansion of the preexisting rule, which generally only permitted a judge to issue warrants for property located inside his or her district,[581] and thereby hindered the government's ability to obtain a remote access search warrant when it could not identify the target's location.[582]

Second, in an investigation of a violation of 18 U.S.C. § 1030(a)(5) (e.g., hacking or malware) that harms computers located in five or more districts, new Rule 41(b)(6)(B) authorizes a judge to issue one warrant to use remote access within or outside the judge's district, and across all of the affected districts. This too is a notable expansion of the preexisting rule, which required law enforcement to submit separate warrant applications in each district where a computer was affected. The change, which will most directly implicate the investigation of multi-district botnet-like schemes, is intended to improve the efficiency and pace of complex computer investigations by allowing a single judge to oversee the investigation.[583]

Over 30 organizations--including Google, the Electronic Frontier Foundation, and the ACLU--submitted written oppositions to the amendments in advance of their adoption.[584] Google noted, for example, that while U.S. law enforcement is generally prohibited from conducting searches in a foreign country, the amendments will allow searches of computers at unknown locations--which could include locations abroad.[585] Google also identified constitutional questions that the amendments are likely to inspire, such as how the government will satisfy the Fourth Amendment's requirement that every warrant "particularly describ[e] the place to be searched."[586] Civil liberties groups also warned that the amendments will lead to "forum shopping" by law enforcement, seeking warrants in districts where judges are most likely to grant them.[587] Congress could have blocked or postponed the amendments, but opposition failed to gain traction on Capitol Hill before the amendments took effect.[588] Remote
access search warrants issued pursuant to the amendments are thus sure to raise new legal questions in the years to come.

IV. International Regulation of Privacy and Data Security

Our separate, additional International Cybersecurity and Data Privacy Outlook and Review addresses international developments of note. Yet again, 2016 saw major developments in the evolution of the data protection and cybersecurity landscape outside the United States:

- The European Union adopted a General Data Protection Regulation governing the processing and transfer of personal data, and stepped up enforcement of data protection regulations in a number of member countries.
- EU and U.S. regulators agreed to a new framework for international data transfers—the Privacy Shield—which has already seen its first legal challenges.
- A number of countries, including Japan and South Korea, amended existing laws, while others, including Argentina, issued new regulations.
- Other countries, for example Mexico, are exploring a regional approach to privacy regulation.

We cover these topics and many more in this year's International Cybersecurity and Data Privacy Outlook and Review.

Of particular note to U.S. companies, on July 12, 2016, the European Commission formally approved the EU-U.S. Privacy Shield ("Privacy Shield"), a framework for navigating the transatlantic transfer of data from the EU to the U.S. The Privacy Shield replaces the EU-U.S. Safe Harbor framework, which was invalidated by the European Court of Justice ("ECJ") on October 6, 2015 in *Maximilian Schrems v. Data Protection Commissioner* (the "Schrems" decision).[589] We provided an in-depth discussion of the Schrems decision in our previous year-end update.[590] In the aftermath of the Schrems decision, EU and U.S. policymakers stepped up their negotiation efforts with respect to a more robust framework to replace the Safe Harbor. The European Commission and the U.S. Government reached a political agreement on the new Privacy Shield framework on February 2, 2016,[591] and published the first draft provisions on February 29, 2016.[592] After receiving a number of responses from important stakeholders, the European Commission included a number of "additional clarifications and improvements," and ultimately approved the framework. Companies can sign up for the Privacy Shield with the U.S. Department of Commerce, which is responsible for verifying that company standards are in compliance with the Privacy Shield.[593]

V. Conclusion

We expect 2017 to be another explosive year in the application and development of privacy and cybersecurity law. Companies and governments will continue to explore the potential uses of personal information. Our public dialogue will continue to evolve with respect to the balance of benefits of big data against concerns for privacy and security. And key entities' technical sophistication (and that of
rivals and adversaries) will certainly continue to develop. We will be tracking these important issues in the year ahead.

[2] Id. at 1549–50.
[5] Id.
[6] Id. at 1545.
[7] Id.
[8] Id. at 1550.
[9] Id. at 1549–50.
[10] Id. at 1549.
[12] Id.
[13] Id.
[14] Id. at 1550.

In re Nickelodeon Consumer Privacy Litig., 827 F. 3d 262, 273-74 (3d Cir. 2016).


Braitberg, 836 F.3d at 930-31.


E.g., Hancock v. Urban Outfitters, 830 F.3d 511, 512 (D.C. Cir. 2016).


phones and minutes is to receive more calls [to receive more calls in violation of the TCPA], thus enabling her to file TCPA lawsuits, she has not suffered an economic injury.


assess with moderate confidence that the group is operating from the Russian Federation and is gathering intelligence on behalf of the Russian government.”).


[41] Id.


[44] In re LinkedIn User Privacy Litig., No. 5:12-cv-03088 (N.D. Cal. Sept. 15, 2015), ECF No. 147 (Order Granting Motion for Settlement).


[52] Id. at 2–3.
[53]  Id. at 6.

[54]  Id. at 10.


[56]  Motion to Dismiss Plaintiffs' Amended Complaint, Torres, No. 6:16-cv-210-PGB-DAB (M.D. Fla. Aug. 19, 2016), ECF No. 74.


[62]  Id.


[64]  Id.


[66]  Although a consolidated complaint is not yet available, one of the early complaints alleging the conduct at issue in In re Sprouts Farmers Market asserts claims for negligence and violations of California consumer protection and unfair business practices statutes. See Amended Complaint, Hernandez v. Sprouts Farmers Market, No. 16-cv-0958-CAB-DHB (S.D. Cal. May 25, 2016), ECF No. 10.


[71] Motion to Compel Arbitration on an Individual Basis, *Varela*, No. 5:16-cv-00577 (C.D. Cal. May 31, 2016), ECF No. 34.

[72] Order re Defendant's Motion to Compel Arbitration or, Alternatively, Motion to Dismiss at 1, 9–10, *Varela*, No. 5:16-cv-00577 (C.D. Cal. July 7, 2016), ECF No. 40.


[74] Motion for Preliminary Injunction re Class Arbitration and Motion to Stay Class Arbitration Pending Appeal, *Varela*, No. 5:16-cv-00577 (C.D. Cal. Nov. 16, 2016), ECF No. 44.


[76] *Varela v. Lamps Plus, Inc.*, No. 16-56085 (9th Cir. docketed July 29, 2016).


[82] *See id.* at 968, n.4.

[83] *Id.* at 1015–16.

[84] *Id.* at 995–96.

[85] *Id.* at 986.

[87] Plaintiffs' Motion for Leave to Proceed Under Pseudonyms, In re Ashley Madison Customer Data Sec. Breach Litig. ("In re Ashley Madison"), No. 4:15-md-02669-JAR (E.D. Mo. Feb. 15, 2016), ECF No. 91.


[89] Avid's Memorandum of Law in Support of Motion to Dismiss Case or Stay and Compel Arbitration at 1, In re Ashley Madison, No. 4:15-md-02669-JAR (E.D. Mo. Aug. 29, 2016), ECF No. 230.


[94] Id. at 1, 30–31.


[99] Order Granting in Part and Denying in Part Motion to Dismiss for Failure to State a Claim, In re Home Depot, No. 1:14-md-02583-TWT (N.D. Ga. May 18, 2016), ECF No. 211.

[100] Motion to Certify Order for Interlocutory Appeal Under § 1292(b), In re Home Depot, No. 1:14-MD-2583-TWT (July 5, 2016), ECF No. 228.

[101] See id. at 2-3.


[104] 819 F.3d 963 (7th Cir. 2016).


[106] Id. at *3.

[107] Lewert, 819 F.3d at 966–67, 969.

[108] 794 F.3d 688 (7th Cir. 2015).

[109] Lewert, 819 F.3d at 967, 969.


[111] Id. at *1.

[112] Id. at *2.

[113] Id. at *1, *6.

[114] Id. at *3.

[115] Id. at *4.


[117] Id. at *7.


[119] Id. at *3.


[122] Order Granting in Part and Denying in Part Motion to Dismiss, *In re Zappos.com, Inc. Customer Data Sec. Breach Litig.* ("In re Zappos"), No. 3:12-cv-00325-RCJ-VPC (D. Nev. May 6, 2016), ECF No. 280. The court also struck the class action language as written, with instructions to "limit the proposed class to individuals who have suffered actual injury." *Id.* at 16.

[123] *Id.* at 7, 10, 11, 14.

[124] *Id.* at 6.


[127] *Id.* at 660.

[128] *Id.*

[129] *Id.* at 661–62.

[130] *Id.* at 662 (quoting *Excavation Techs., Inc. v. Columbia Gas Co. of Pa.*, 985 A.2d 840, 841 (Pa. 2009)).


[132] *Id.* at *7 (internal quotation marks omitted) (citing *Busse v. Motorola Inc.*, 813 N.E.2d 1013, 1017 (Ill. App. Ct. 2004)).


[134] *Id.*

Id. at 11-12.


Id. ¶¶ 25, 35.

Memorandum and Order Granting Motion for Final Approval of Consumer Settlement and Motion for Payment of Service Awards and Attorney's Fees and Expenses, In re Target, No. 0:14-md-02522-PAM (D. Minn. Nov. 17, 2015), ECF No. 645.

Id. at 1-2, 8; Consumer Settlement Agreement and Release ¶¶ 1.2.1, 5.1–5.2.4, In re Target, No. 0:14-md-02522-PAM (D. Minn. Mar. 18, 2015), ECF No. 358-1.

Consumer Settlement Agreement and Release ¶ 5.2–5.2.4, In re Target, No. 0:14-md-02522-PAM (D. Minn. Mar. 18, 2015), ECF No. 358-1.

See Consumer Plaintiffs' Memorandum in Support of Motion for Appeal Bond at 2, In re Target, No. 0:14-md-02522-PAM (D. Minn. Dec. 21, 2015), ECF No. 680 (documenting the appeals of objectors Sciaroni, Olson, Gibson, and non-class member Miorelli).


Olson v. Target Corp., No. 15-3912 (8th Cir. Dec. 21, 2015).

Memorandum and Order Granting Motion for Final Approval of Financial Institutions' Class Action Settlement and Motion for Attorneys' Fees and Expenses and Service Payments, *In re Target*, No. 0:14-md-02522-PAM (D. Minn. May 12, 2016), ECF No. 758.

Financial Institutions' Settlement Agreement and Release ¶ 1.41, *In re Target*, No. 0:14-md-02522-PAM (D. Minn. Dec. 2, 2015), ECF No. 653-1. MasterCard's Account Data Compromise program is a program for assessment of fraud recovery, whereby merchant banks (or their merchants, e.g., Target) reimburse MasterCard's issuing banks for fraud and expenses resulting from data breaches. Target also separately settled claims by Visa and certain of Visa's issuing banks for $63.5 million. See id. at 5.

Memorandum and Order Granting Motion for Final Approval of Financial Institutions' Class Action Settlement and Motion for Attorneys' Fees and Expenses and Service Payments at 4, 15, *In re Target*, No. 0:14-md-02522-PAM (D. Minn. May 12, 2016), ECF No. 758.


Mem. and Order Granting Mot. for Final Approval of Financial Institutions' Class Action Settlement and Mot. for Att'y Fees and Expenses and Service Payments, *In re Target*, No. 0:14-md-02522-PAM (D. Minn. May 12, 2016), ECF No. 758 (adopting Settlement Agreement, ECF No. 653-1).


Mem. and Order Granting Mot. for Final Approval of Consumer Settlement and Mot for Payment of Service Awards and Fees and Expenses, *In re Target*, No. 0:14-md-02522-PAM (D. Minn. Nov. 16, 2016), ECF No. 645 (approving Settlement Agreement, ECF No. 358-1).


Id. at 11-12.

Id. at 13-14.

Id. at 14.

Id. at 14-18. The Court arrived at similar conclusion in considering plaintiffs’ claims for corporate waste and violations of Section 14(a) of the Securities Exchange Act. Id. at 22, 30.

Id. at 22.

Id. at 30.


Id. at 284–85.

Id. at 282.

Id. at 295.


Id.

Id.

[201] *Id.* at *16 ("[I]t appears that there is no 'real and immediate threat of repeated injury in the future.'").


[205] *Id.* at 17.

[206] *Id.* at 5.

[207] *Id.* at 6.

[208] *Id.*


[210] *Id.*


[214] *See, e.g.,* Saulsberry, 2016 WL 3456939, at *14–16 (concluding that the section 632.7 class satisfied the commonality requirement for class certification while the section 632 class did not, although ultimately denying certification of the section 632.7 class because the class representative was not typical of the class).


[216] *Id.*

[218]  *Id.* at 2, 6.


[222]  *Id.* at 2.


[228]  *Id.*


[231]  *Id.*


Satchell v. Sonic Notify, Inc., No. 4:16-cv-04961 (N.D. Cal).

Id., ECF No. 28.


833 F.3d 619, 623 (6th Cir. 2016).

Id. at 624.

Id. at 634.

Id. at 633.

47 U.S.C. §§ 227 et seq.


Id. at 7978 ¶ 24 (2015).

Id. at 7989–90 ¶ 47.


Compare Konopca v. Comcast Corp., No. 15-6044 (FLW)(DEA), 2016 WL 1645157 (D.N.J. Apr. 26, 2016) (denying motion to stay pending the D.C. Circuit's ruling on the July 2015 omnibus order because lengthy delay could potentially harm the plaintiff and any ruling on the appeal would not be dispositive for the matter before the court), and Schwyhart v. AmSher Collection Servs., Inc., No. 2:15-CV-01175-JEO, 2016 WL 1620096 (N.D. Ala. Apr. 22, 2016) (denying motion to stay pending the D.C. Circuit's ruling on the July 2015 omnibus order because of the possibility of an indefinite delay in the appeals process), with Rose v. Wells Fargo Advisors, LLC, No. 1:16-CV-562-CAP, 2016 WL 3369283 (N.D. Ga. June 14, 2016) (granting motion to stay pending the D.C. Circuit's ruling on the July 2015 omnibus order because potential prejudice to the plaintiff was minimal and "the defendant may suffer hardship in conducting discovery and trial preparation in light of the uncertain difference between 'potential' capacity and 'theoretical' capacity under the definition of an [autodialer]," which the D.C. Circuit may clarify).

136 S. Ct. 1540 (2016), as revised (May 24, 2016).

Id. at 1549.

[251] See, e.g., Romero v. Dep't Stores Nat'l Bank, No. 15-CV-193-CAB-MDD, 2016 WL 4184099, at *4 (S.D. Cal. Aug. 5, 2016) (finding the "[p]laintiff's failure to connect any of these claimed injuries in fact [of lost time, aggravation, and distress] with any (or each) specific TCPA violation is alone fatal to Plaintiff's standing argument"); Stoops v. Wells Fargo Bank, N.A., No. CV 3:15-83, 2016 WL 3566266, at *12 (W.D. Pa. June 24, 2016) ("Because Plaintiff has admitted that her only purpose in purchasing her cell phones and minutes is to receive more calls, thus enabling her to file TCPA lawsuits, she has not suffered an economic injury.").


[253] Id.

[254] Preliminary Approval Order, Aranda, ECF No. 505; see also Minute Entry, Aranda, ECF No. 512 (rescheduling final approval hearing for February 23, 2017).


[257] Id.


[259] 803 F.3d 1251, 1253, 1257 (11th Cir. 2015).

[260] Another unpublished decision in 2015 similarly concluded that a "subscriber" must be required to exchange "money and/or personal information in order to receive a future and recurrent benefit." Austin-Spearman v. AMC Network Ent., LLC, 98 F. Supp. 3d 662, 669 (S.D.N.Y. Apr. 7, 2015).

In addition to affirming the dismissal of the VPPA claims against Viacom on the aforementioned grounds, the appeals court also affirmed the dismissal of VPPA claims against Google because it was not a "video tape service provider." The VPPA defines a "video tape service provider," in part, as any person "engaged in the business . . . of rental, sale, or delivery of prerecorded video cassette tapes or similar audio visual materials." 18 U.S.C. § 2710(a)(4). In reaching its conclusion, the appeals court noted that Google could not be liable for receipt of "personally identifiable information" because the VPPA only creates liability for disclosure of such information. 827 F.3d 262, 281 (3d Cir. 2016). The Supreme Court denied Plaintiffs' petition for writ of certiorari, letting stand the decision to dismiss the VPPA claims, and maintaining the split with the First Circuit's decision in Gannett. See C.A.F. et al. v. Viacom Inc. et al., No. 16-346 (U.S. Jan. 9, 2017).

Also in line with the Third and Eleventh Circuits, the Michigan Supreme Court unanimously held that a Pandora user was not a "customer" under Michigan's state analog to the VPPA because he neither rented nor borrowed anything from Pandora. In re Certified Question from the U.S. Court of Appeals for the Ninth Circuit, Deacon v. Pandora Media Inc., No. 151104 (Mich. Sup. 2016).

The court cited Yershov, 2016 WL 4607868, at *7; In re Nickelodeon, 827 F.3d at 274; Sterk v. Redbox Automated Retail, LLC, 770 F.3d 618, 623 (7th Cir. 2014); and Austin–Spearman v. AMC Network Entm't LLC, 98 F. Supp. 3d 662, 668 (S.D.N.Y. 2015).

[276] Harrold v. Levi Strauss & Co., 236 Cal. App. 4th 1259, 1265 (2015); see also Davis v. Devanlay Retail Grp., 785 F.3d 359 (9th Cir.), ECF No. 41 (in declining to respond to a question certified by the Ninth Circuit but addressed in Harrold, the California Supreme Court expressly situated Harrold as controlling precedent).


[281] See Medellin v. Ikea U.S. West Incorp., No. 15-55174 (9th Cir.) (Dkt. No. 43).


[285] Id.


[288] Id. at 353-54.


[290] Alaska, see H.B. 96, 29th Leg. (Ak. 2015), and Idaho, see H.B. 511, 2014 Reg. Sess. (Id. 2014), introduced laws to protect biometric data in 2015 and 2014, respectively, while an amendment that would include biometrics as personal information in California failed passage in August 2016, see A.B. 83, 2015 Reg. Sess. (Ca. 2015). Washington passed a bill defining biometric identifiers and regulating their
use for commercial purposes in February 2016, though it was returned to the House Rules Committee in March. See H.B. 1094, 2015 Reg. Sess. (Wa. 2015).

[291] 740 ILCS § 14/1 et seq.

[292] Texas's law is similar to BIPA; however, it has not been litigated because it does not provide for a private right of action.

[293] 740 ILCS §§ 14/15(b)-(e).

[294] Id. § 14/20.


[296] See, e.g., Mot. to Dismiss, In re Facebook Biometric, ECF No. 69 (N.D. Cal. Oct. 9, 2015); Mot. to Dismiss, Rivera, ECF No. 33 (N.D. Ill. May 18, 2016); Mot. to Dismiss, Norberg, ECF No. 26 (N.D. Ill. July 31, 2015). See Siegal v. Snapchat was voluntarily dismissed with no substantive briefing.


[302] Id. at *3–4.


GIBSON DUNN


[310] ASUSTeK Decision and Order at 7.


[312] Id. at 5–6, 8, 11–13.

[313] Id. at 6–7.

[314] Id. at 10–13.


[318] See Appellants' Reply Brief, Cahen v. Toyota Motor Corp., No. 16-15496 (9th Cir. Nov. 9, 2016).


[320] Id.

[321] Id. at *2.

[322] Id. at *3–4.


[326] *Id.*

[327] *Id.*


[330] *Id.*

[331] *Id.*


Id.


See National Association of Insurance Commissioners, Cybersecurity (Nov. 17, 2016), available at http://www.naic.org/cipr_topics/topic_cyber_risk.htm ("Most businesses are familiar with their commercial insurance policies providing general liability coverage to protect the business from injury or property damage. However, most standard commercial lines policies do not cover many of the cyber risks mentioned above. To cover these unique cyber risks through insurance requires the purchase of a special cyber liability policy.").


Id. at *2.

Id.

Id. at *9.

Shaun Waterman, Experts: Cyber insurance market full of "trap doors," Cyberscoop (Oct. 25, 2016), available at https://www.cyberscoop.com/experts-cyber-insurance-market-full-trap-doors/ ("Cyber insurance offered by major carriers is extremely diverse, with the absence of standardization meaning the market is full of 'trapdoors' for the unwary buyer . . . .").

799 F.3d 236 (3d Cir. 2015).


Final Order, supra note 352.

Id. at 2–3.

Id. at 3–4.


Id.


Id.


Id. at 6–7.


[366] Id.

[367] Id.

[368] Id.


[375] *F.T.C. v. AT&T Mobility LLC*, No. 3:14-cv-04785-EMC (9th Cir. 2016).

[376] See id. at 6.

[377] Id. at 9.


Id.

Id.; see also Department of Health and Human Services, HIPAA Privacy, Security, and Breach Notification Audit Program (no date), available at https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/audit/index.html.


Resolution Agreement, Department of Health and Human Services (July 8, 2016), available at https://www.hhs.gov/sites/default/files/Advocate_racap.pdf.

Id. at 2.

Press Release, Department of Health and Human Services, supra note 383.


Lambert & Barlyn, supra note 389.

Id.
"In general, the priorities reflect certain practices and products that OCIE perceives to present potentially heightened risk to investors and/or the integrity of the U.S. capital markets." SEC, Examination Priorities for 2016 (Jan. 11, 2016), available at https://www.sec.gov/about/offices/ocie/national-examination-program-priorities-2016.pdf; see also SEC, Examination Priorities for 2017 (Jan. 12, 2017), available at https://www.sec.gov/about/offices/ocie/national-examination-program-priorities-2017.pdf.

Id. at 3.

Id.


Id.

OCIE, supra note 395, at 3.


Id.

Id.


Morgan Stanley Smith Barney LLC, Exchange Act Release No. 78021, Investment Advisers Act Release No. 4415, Administrative Proceeding File No. 3-17280, at 2 (June 8, 2016). The Safeguards Rule requires broker-dealers and investment advisers to adopt policies and procedures reasonably designed to: (1) insure the security and confidentiality of customer records and information; (2) protect against any anticipated threats or hazards to the security or integrity of customer records and information; and (3) protect against unauthorized access to or use of customer records or information that could result in substantial harm or inconvenience to any customer. Id. at 5.
[404] *Id.* at 6.

[405] *Id.* at 4.

[406] *Id.*

[407] *Id.* at 5.

[408] *Id.*

[409] In both R.T. Jones and Morgan Stanley, the SEC did not allege that improperly accessed personal data had caused harm to customers and/or clients. *See supra* note 401 (finding no evidence that R.T. Jones' customers were financially harmed by the data breach).


[411] *Id.* ¶ 166. Providers must also obtain affirmative opt-in consent before making any material retroactive changes to the use of any personally-identifiable information, including sensitive and non-sensitive information. *Id.* ¶ 195.

[412] *Id.* ¶ 177. The new rules are also applicable to voice services, and call-detail record information is considered "sensitive" information in that context. *See id.* ¶ 167.

[413] *Id.* ¶ 199.

[414] *Id.* ¶ 201.


[418] Edge providers may be subject to other federal and state laws and regulations governing their collection and use of consumer information, as well as to limitations in consumer agreements with their users. *See Dissenting Statement of Commissioner Ajit Pai*, dated Oct. 27, 2016.


[421] Id.

[422] Id.


[424] Id. at 2.

[425] Id.

[426] Id. at 2.

[427] Id.

[428] Id. at 25.


[430] Id. at 5.

[431] Id. at 12.

[432] See Christopher G. Cwalina, et al., CFPB Expands UDAAP Jurisdiction in First Foray into Data Security Enforcement, 133 Banking L.J. 279 (Mar. 8, 2016) ("The Dodd-Frank Act excludes from the definition of enumerated consumer laws placed under the CFPB's jurisdiction the key provisions of the Gramm-Leach-Bliley Act, the primary federal law regulating data security. The CFPB's consent order with Dwolla demonstrates that the CFPB has gotten around this limitation by self-defining its Unfair, Deceptive or Abusive Acts and Practices ("UDAAP") authority as encompassing data security matters.").


[434] Id. at 4, 15-16.


[442] See id.; supra note 440.


[444] Id.

[445] Id. (The FTC’s safe harbor program allows an organization to create and comply with self-regulatory guidelines, which the Commission must approve as equivalent the protections provided under


Supra note 453.

Id. at 11. Originally, regulated entities were to have 180 days from January 1, 2017 to become compliant with all provisions. Now, compliance within 180 days of March 1, 2017 would be required for certain provisions, including mandates to: develop a cybersecurity program; implement written cybersecurity policies approved by either a senior officer or the Board of Directors; review access privileges; designate a Chief Information Security Officer who reports to the Board of Directors; develop a cybersecurity incident response plan; and begin making required reports of security incidents to DFS, as discussed more fully below. Regulated entities now would have one year from March 1, 2017 to take additional steps, including: perform a risk assessment of their information systems; implement multi-factor authentication for external access to internal databases and privileged access to nonpublic information; and begin cybersecurity training for all personnel. In addition, regulated entities now would have eighteen months from March 1, 2017 to implement a third set of requirements: develop audit trail systems to track and maintain data to allow for reconstruction of all financial transactions and accounting necessary to detect and respond to a cybersecurity event; establish a data retention policy to ensure the secure disposal of nonpublic data; develop written policies and procedures to ensure the security of in-house developed applications; and establish policies and procedures to monitor the activity of authorized users, detect unauthorized access, and encrypt (or impose alternative controls to protect) nonpublic information. Finally, regulated entities now would have two years from March 1, 2017 to establish written policies and procedures to ensure the security of data that is accessible to or held by third-party service providers.

Id. at 5

Id. at 10.

Id. at 2. The original proposed rules defined "Nonpublic information" to include "any information that can be used to distinguish or trace an individual's identity." The revised definition is narrowed to include "Any information concerning an individual which because of name, number personal mark, or other identifier can be used to identify such individual, in combination with any one or more of the following data elements: (i) social security number, (ii) drivers' license number or non-driver identification card number, (iii) account number, credit or debit card number, (iv) any security code, access code or password that would permit access to an individual's financial account; or (v) biometric records."


Id.
The Commission's goals also include, "ensuring public safety and economic and national security, fostering discovery and development of new technical solutions, and bolstering partnerships between Federal, State, and local government and the private sector in the development, promotion, and use of cybersecurity technologies, policies, and best practices." See Executive Order, supra note 462.


See id., for a full list of Commission members.


Id. at 11.

Id. at 43.

Id. at 47.

Press Release, supra note 469.


Id.

[478] Id.


[480] Id.

[481] Id.


[483] Id.


[485] Id.

[486] Id.


[488] See supra note 484.


[490] Id. § 2.

[491] Id.

[492] Id. at Preamble.

[494] *Id.* ("The tech industry is concerned about the broader consequences of not addressing some of the international community's post-Snowden concerns.").


[497] *Id.*


[499] *Id.* § 2.

[500] *Id.*


[503] *Id.* § 3.


[505] *See id.* (providing a list of size standards based on a company's North American Industry Classification System ("NAICS") code).


[510] *Id.* § 2.

[511] *Id.*

[512] Lennon, supra note 508.


[522] *See id.*


[526] See id.


[528] See An Act to Amend and Reenact §§ 2.2-3701, 2.2-3704, 2.2-3705.1 through 2.2-3705.7, 2.2-3711, and 2.2-3713 of the Code of Virginia and to Amend the Code of Virginia by Adding a Section Numbered 2.2-3704.01, Relating to the Virginia Freedom of Information Act, H.B. 817, Reg. Sess. 2016 (Va. 2016).


[531] See id.

[532] See id.

[533] Id.


[537] Id.

[538] Id.

[539] See supra note 514.

[540] See id.


[544] Specifically, a court must grant a government application for a nondisclosure order if it finds reason to believe that disclosure will result in: (1) Endangering the life or physical safety of an individual; (2) Flight from prosecution; (3) Destruction or tampering with evidence; (4) Intimidation of potential witnesses; or (5) Otherwise seriously jeopardizing an investigation or unduly delaying a trial. 18 U.S.C. § 2705(b).

[545] Id.


[547] See id. at ¶ 24–25.


[552] Id. at 22–25.

[553] Id. at 20-21.

[554] Id. at 21.

[555] Id. at 22–22.

[556] Id. at 26–27.

[557] Id. at 9–12.


[570]  Id. at 472.
Matter of Warrant to Search a Certain E–Mail Account Controlled & Maintained by Microsoft Corp., 829 F.3d 197, 200 (2d Cir. 2016).

Id. at 210 (quoting Morrison v. Nat'l Australia Bank Ltd., 561 U.S. 247, 255 (2010)).

Id. at 211-13.

Id. at 219-20.

Id. at 212 (quoting In re Terrorist Bombings of U.S. Embassies in E. Africa, 552 F.3d 157, 169 (2d Cir. 2008)).

Id. at 220.

Id. at 222.

Jonathan Stempel, Microsoft wins landmark appeal over seizure of foreign emails, REUTERS (July 14, 2016), available at http://www.reuters.com/article/us-microsoft-usa-warrant-idUSKCN0ZU1RJ.


Id. at 6–7.

Id.


[586] Id. at 8-9.


Gibson Dunn's lawyers are available to assist with any questions you may have regarding these issues. For further information, please contact the Gibson Dunn lawyer with whom you usually work or any of the following leaders and members of the firm's Privacy, Cybersecurity and Consumer Protection Group. In addition, Gibson Dunn is pleased to announce that our EU expertise has been substantially augmented with the arrival of four technology lawyers, anchored by partner Ahmed Baladi, to our Paris Office. Mr. Baladi and his team bring additional expertise with EU data privacy and cybersecurity, information technology and digital transactions, and outsourcing.

United States
Alexander H. Southwell - Chair, New York (+1 212-351-3981, asouthwell@gibsondunn.com)
M. Sean Royall - Dallas (+1 214-698-3256, sroyall@gibsondunn.com)
Debra Wong Yang - Los Angeles (+1 213-229-7472, dwongyang@gibsondunn.com)
Howard S. Hogan - Washington, D.C. (+1 202-887-3640, hhogan@gibsondunn.com)
Joshua A. Jessen - Orange County/Palo Alto (+1 949-451-4114/+1 650-849-5375, jjessen@gibsondunn.com)
Shaalu Mehra - Palo Alto (+1 650-849-5282, smehra@gibsondunn.com)
Karl G. Nelson - Dallas (+1 214-698-3203, knelson@gibsondunn.com)
Eric D. Vandevelde - Los Angeles (+1 213-229-7186, evandevelde@gibsondunn.com)
Benjamin B. Wagner - Palo Alto (+1 650-849-5395, bwagner@gibsondunn.com)
Michael Li-Ming Wong - San Francisco/Palo Alto (+1 415-393-8333/+1 650-849-5393, mwong@gibsondunn.com)
Ryan T. Bergsieker - Denver (+1 303-298-5774, rbergsieker@gibsondunn.com)
Richard H. Cunningham - Denver (+1 303-298-5752, rhcunningham@gibsondunn.com)
Jeana Bisnar Maute - Palo Alto (+1 650-849-5348, jbisnarmaute@gibsondunn.com)

Europe
Ahmed Baladi - Paris (+33 (0)1 56 43 13 00, abaladi@gibsondunn.com)
James A. Cox - London (+44 (0)20 7071 4250, jaco@gibsondunn.com)
Patrick Doris - London (+44 (0)20 7071 4276, pdoris@gibsondunn.com)
Andrés Font Galarza - Brussels (+32 2 554 7230, afontgalarza@gibsondunn.com)
Bernard Grinspan - Paris (+33 (0)1 56 43 13 00, bgri@gibsondunn.com)
Penny Madden - London (+44 (0)20 7071 4226, pmadden@gibsondunn.com)
Jean-Philippe Robé - Paris (+33 (0)1 56 43 13 00, jrobe@gibsondunn.com)
Michael Walther - Munich (+49 89 189 33-180, mwalther@gibsondunn.com)
Nicolas Autet - Paris (+33 (0)1 56 43 13 00, nautet@gibsondunn.com)
Kat Gesing - Munich (+49 89 189 33-180, kgesing@gibsondunn.com)
Sarah Wazen - London (+44 (0)20 7071 4203, swazen@gibsondunn.com)
Emmanuelle Bartoli - Paris (+33 (0)1 56 43 13 57, ebartoli@gibsondunn.com)
Eryk L. Dziadykiewicz - Brussels (+32 2 554 72 03, edziadykiewicz@gibsondunn.com)
Alejandro Guerrero Perez - Brussels (+32 2 554 7218, aguerrerperez@gibsondunn.com)
Questions about SEC disclosure issues concerning data privacy and cybersecurity can also be addressed to the following leaders and members of the Securities Regulation and Corporate Disclosure Group:

James J. Moloney - Orange County, CA (+1 949-451-4343, jmoloney@gibsondunn.com)
Elizabeth Ising - Washington, D.C. (+1 202-955-8287, eising@gibsondunn.com)
Lori Zyskowski - New York (+1 212-351-2309, lzyskowski@gibsondunn.com)
Panel II:

Artificial Intelligence: Issues of Technology, Ethics, and Law

Moderator:
Lala Qadir

Discussants:
Gregory Neal Akers
Michael Page
Lt. Col. Alan Schuller
David Vladeck
PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE

Executive Office of the President
National Science and Technology Council
Committee on Technology

October 2016
About the National Science and Technology Council

The National Science and Technology Council (NSTC) is the principal means by which the Executive Branch coordinates science and technology policy across the diverse entities that make up the Federal research and development (R&D) enterprise. One of the NSTC’s primary objectives is establishing clear national goals for Federal science and technology investments. The NSTC prepares R&D packages aimed at accomplishing multiple national goals. The NSTC’s work is organized under five committees: Environment, Natural Resources, and Sustainability; Homeland and National Security; Science, Technology, Engineering, and Mathematics (STEM) Education; Science; and Technology. Each of these committees oversees subcommittees and working groups that are focused on different aspects of science and technology. More information is available at www.whitehouse.gov/ostp/nstc.

About the Office of Science and Technology Policy

The Office of Science and Technology Policy (OSTP) was established by the National Science and Technology Policy, Organization, and Priorities Act of 1976. OSTP’s responsibilities include advising the President in policy formulation and budget development on questions in which science and technology are important elements; articulating the President’s science and technology policy and programs; and fostering strong partnerships among Federal, state, and local governments, and the scientific communities in industry and academia. The Director of OSTP also serves as Assistant to the President for Science and Technology and manages the NSTC. More information is available at www.whitehouse.gov/ostp.

Acknowledgments

This document was developed through the contributions of staff from OSTP, other components of the Executive Office of the President, and other departments and agencies. A special thanks and appreciation to everyone who contributed.

Copyright Information

This is a work of the U.S. Government and is in the public domain. It may be freely distributed, copied, and translated; acknowledgment of publication by the Office of Science and Technology Policy is appreciated. Any translation should include a disclaimer that the accuracy of the translation is the responsibility of the translator and not OSTP. It is requested that a copy of any translation be sent to OSTP. This work is available for worldwide use and reuse and under the Creative Commons CC0 1.0 Universal license.
Dear colleagues:

Advances in Artificial Intelligence (AI) technology have opened up new markets and new opportunities for progress in critical areas such as health, education, energy, and the environment. In recent years, machines have surpassed humans in the performance of certain specific tasks, such as some aspects of image recognition. Experts forecast that rapid progress in the field of specialized artificial intelligence will continue. Although it is very unlikely that machines will exhibit broadly-applicable intelligence comparable to or exceeding that of humans in the next 20 years, it is to be expected that machines will reach and exceed human performance on more and more tasks.

As a contribution toward preparing the United States for a future in which AI plays a growing role, this report surveys the current state of AI, its existing and potential applications, and the questions that are raised for society and public policy by progress in AI. The report also makes recommendations for specific further actions by Federal agencies and other actors. A companion document lays out a strategic plan for Federally-funded research and development in AI. Additionally, in the coming months, the Administration will release a follow-on report exploring in greater depth the effect of AI-driven automation on jobs and the economy.

The report was developed by the NSTC’s Subcommittee on Machine Learning and Artificial Intelligence, which was chartered in May 2016 to foster interagency coordination, to provide technical and policy advice on topics related to AI, and to monitor the development of AI technologies across industry, the research community, and the Federal Government. The report was reviewed by the NSTC Committee on Technology, which concurred with its contents. The report follows a series of public-outreach activities spearheaded by the White House Office of Science and Technology Policy (OSTP) in 2016, which included five public workshops co-hosted with universities and other associations that are referenced in this report.

OSTP also published a Request for Information (RFI) in June 2016, which received 161 responses. The submitted comments were published by OSTP on September 6, 2016. Consistent with the role of Big Data as an enabler of AI, this report builds on three previous Administration reports on Big Data referenced in this report.

In the coming years, AI will continue to contribute to economic growth and will be a valuable tool for improving the world, as long as industry, civil society, and government work together to develop the positive aspects of the technology, manage its risks and challenges, and ensure that everyone has the opportunity to help in building an AI-enhanced society and to participate in its benefits.

Sincerely,

John P. Holdren
Assistant to the President for Science and Technology
Director, Office of Science and Technology Policy

Megan Smith
U.S. Chief Technology Officer
The following Federal departments and agencies are represented on the Subcommittee on Machine Learning and Artificial Intelligence and through it, work together to monitor the state of the art in machine learning (ML) and AI (within the Federal Government, in the private sector, and internationally), to watch for the arrival of important technology milestones in the development of AI, to coordinate the use of and foster the sharing of knowledge and best practices about ML and AI by the Federal Government, and to consult in the development of Federal research and development priorities in AI:

- Department of Commerce (Co-Chair)
- Department of Defense
- Department of Education
- Department of Energy
- Department of Health and Human Services
- Department of Homeland Security
- Department of Justice
- Department of Labor
- Department of State
- Department of Transportation
- Department of Treasury
- Department of Veterans Affairs
- United States Agency for International Development
- Central Intelligence Agency
- General Services Administration
- National Science Foundation
- National Security Agency
- National Aeronautics and Space Administration
- Office of the Director of National Intelligence
- Social Security Administration

The following offices of the Executive Office of the President are also represented on the Subcommittee:

- Council of Economic Advisers
- Domestic Policy Council
- Office of Management and Budget
- Office of Science and Technology Policy (Co-Chair)
- Office of the Vice President
- National Economic Council
- National Security Council
Contents

Executive Summary .................................................................................................................. 1

Introduction .......................................................................................................................... 5
  A Brief History of AI .......................................................................................................... 5
  What is Artificial Intelligence? .......................................................................................... 6
  The Current State of AI .................................................................................................... 7

Public Outreach and Development of this Report ............................................................... 12

Applications of AI for Public Good .................................................................................... 13

AI in the Federal Government ............................................................................................ 15

AI and Regulation ............................................................................................................... 17
  Case Study: Autonomous Vehicles and Aircraft ............................................................ 18

Research and Workforce .................................................................................................... 23
  Monitoring Progress in AI ............................................................................................... 23
  Federal Support for AI Research ..................................................................................... 25
  Workforce Development and Diversity .......................................................................... 26

AI, Automation, and the Economy ....................................................................................... 29

Fairness, Safety, and Governance ....................................................................................... 30
  Justice, Fairness, and Accountability .............................................................................. 30
  Safety and Control ........................................................................................................... 32

Global Considerations and Security .................................................................................. 35
  International Cooperation ............................................................................................... 35
  AI and Cybersecurity ....................................................................................................... 36
  AI in Weapon Systems ...................................................................................................... 37

Conclusion ........................................................................................................................... 39

Recommendations in this Report ....................................................................................... 40

Acronyms ............................................................................................................................ 43

References ........................................................................................................................... 45
Executive Summary

As a contribution toward preparing the United States for a future in which Artificial Intelligence (AI) plays a growing role, we survey the current state of AI, its existing and potential applications, and the questions that are raised for society and public policy by progress in AI. We also make recommendations for specific further actions by Federal agencies and other actors. A companion document called the National Artificial Intelligence Research and Development Strategic Plan lays out a strategic plan for Federally-funded research and development in AI.

Applications of AI for Public Good

One area of great optimism about AI and machine learning is their potential to improve people’s lives by helping to solve some of the world’s greatest challenges and inefficiencies. Many have compared the promise of AI to the transformative impacts of advancements in mobile computing. Public- and private-sector investments in basic and applied R&D on AI have already begun reaping major benefits to the public in fields as diverse as health care, transportation, the environment, criminal justice, and economic inclusion. The effectiveness of government itself is being increased as agencies build their capacity to use AI to carry out their missions more quickly, responsively, and efficiently.

AI and Regulation

AI has applications in many products, such as cars and aircraft, which are subject to regulation designed to protect the public from harm and ensure fairness in economic competition. How will the incorporation of AI into these products affect the relevant regulatory approaches? In general, the approach to regulation of AI-enabled products to protect public safety should be informed by assessment of the aspects of risk that the addition of AI may reduce alongside the aspects of risk that it may increase. If a risk falls within the bounds of an existing regulatory regime, moreover, the policy discussion should start by considering whether the existing regulations already adequately address the risk, or whether they need to be adapted to the addition of AI. Also, where regulatory responses to the addition of AI threaten to increase the cost of compliance, or slow the development or adoption of beneficial innovations, policymakers should consider how those responses could be adjusted to lower costs and barriers to innovation without adversely impacting safety or market fairness.

Currently relevant examples of the regulatory challenges that AI-enabled products present are found in the cases of automated vehicles (AVs, such as self-driving cars) and AI-equipped unmanned aircraft systems (UAS, or “drones”). In the long run, AVs will likely save many lives by reducing driver error and increasing personal mobility, and UAS will offer many economic benefits. Yet public safety must be protected as these technologies are tested and begin to mature. The Department of Transportation (DOT) is using an approach to evolving the relevant regulations that is based on building expertise in the Department, creating safe spaces and test beds for experimentation, and working with industry and civil society to evolve performance-based regulations that will enable more uses as evidence of safe operation accumulates.

Research and Workforce

Government also has an important role to play in the advancement of AI through research and development and the growth of a skilled, diverse workforce. A separate strategic plan for Federally-funded AI research and development is being released in conjunction with this report. The plan discusses the role of Federal R&D, identifies areas of opportunity, and recommends ways to coordinate R&D to maximize benefit and build a highly-trained workforce.
Given the strategic importance of AI, moreover, it is appropriate for the Federal Government to monitor developments in the field worldwide in order to get early warning of important changes arising elsewhere in case these require changes in U.S. policy.

The rapid growth of AI has dramatically increased the need for people with relevant skills to support and advance the field. An AI-enabled world demands a data-literate citizenry that is able to read, use, interpret, and communicate about data, and participate in policy debates about matters affected by AI. AI knowledge and education are increasingly emphasized in Federal Science, Technology, Engineering, and Mathematics (STEM) education programs. AI education is also a component of Computer Science for All, the President’s initiative to empower all American students from kindergarten through high school to learn computer science and be equipped with the computational thinking skills they need in a technology-driven world.

**Economic Impacts of AI**

AI’s central economic effect in the short term will be the automation of tasks that could not be automated before. This will likely increase productivity and create wealth, but it may also affect particular types of jobs in different ways, reducing demand for certain skills that can be automated while increasing demand for other skills that are complementary to AI. Analysis by the White House Council of Economic Advisors (CEA) suggests that the negative effect of automation will be greatest on lower-wage jobs, and that there is a risk that AI-driven automation will increase the wage gap between less-educated and more-educated workers, potentially increasing economic inequality. Public policy can address these risks, ensuring that workers are retrained and able to succeed in occupations that are complementary to, rather than competing with, automation. Public policy can also ensure that the economic benefits created by AI are shared broadly, and assure that AI responsibly ushers in a new age in the global economy.

**Fairness, Safety, and Governance**

As AI technologies move toward broader deployment, technical experts, policy analysts, and ethicists have raised concerns about unintended consequences of widespread adoption. Use of AI to make consequential decisions about people, often replacing decisions made by human-driven bureaucratic processes, leads to concerns about how to ensure justice, fairness, and accountability—the same concerns voiced previously in the Administration’s *Big Data: Seizing Opportunities, Preserving Values* report of 2014,¹ as well as the Report to the President on *Big Data and Privacy: A Technological Perspective* published by the President’s Council of Advisors on Science and Technology in 2014.² Transparency concerns focus not only on the data and algorithms involved, but also on the potential to have some form of explanation for any AI-based determination. Yet AI experts have cautioned that there are inherent challenges in trying to understand and predict the behavior of advanced AI systems.

Use of AI to control physical-world equipment leads to concerns about safety, especially as systems are exposed to the full complexity of the human environment. A major challenge in AI safety is building systems that can safely transition from the “closed world” of the laboratory into the outside “open world” where unpredictable things can happen. Adapting gracefully to unforeseen situations is difficult yet necessary for safe operation. Experience in building other types of safety-critical systems and infrastructure, such as aircraft, power plants, bridges, and vehicles, has much to teach AI practitioners

---


about verification and validation, how to build a safety case for a technology, how to manage risk, and how to communicate with stakeholders about risk.

At a technical level, the challenges of fairness and safety are related. In both cases, practitioners strive to avoid unintended behavior, and to generate the evidence needed to give stakeholders justified confidence that unintended failures are unlikely.

Ethical training for AI practitioners and students is a necessary part of the solution. Ideally, every student learning AI, computer science, or data science would be exposed to curriculum and discussion on related ethics and security topics. However, ethics alone is not sufficient. Ethics can help practitioners understand their responsibilities to all stakeholders, but ethical training should be augmented with technical tools and methods for putting good intentions into practice by doing the technical work needed to prevent unacceptable outcomes.

**Global Considerations and Security**

AI poses policy questions across a range of areas in international relations and security. AI has been a topic of interest in recent international discussions as countries, multilateral institutions, and other stakeholders have begun to access the benefits and challenges of AI. Dialogue and cooperation between these entities could help advance AI R&D and harness AI for good, while also addressing shared challenges.

Today’s AI has important applications in cybersecurity, and is expected to play an increasing role for both defensive and offensive cyber measures. Currently, designing and operating secure systems requires significant time and attention from experts. Automating this expert work partially or entirely may increase security across a much broader range of systems and applications at dramatically lower cost, and could increase the agility of the Nation’s cyber-defenses. Using AI may help maintain the rapid response required to detect and react to the landscape of evolving threats.

Challenging issues are raised by the potential use of AI in weapon systems. The United States has incorporated autonomy in certain weapon systems for decades, allowing for greater precision in the use of weapons and safer, more humane military operations. Nonetheless, moving away from direct human control of weapon systems involves some risks and can raise legal and ethical questions.

The key to incorporating autonomous and semi-autonomous weapon systems into American defense planning is to ensure that U.S. Government entities are always acting in accordance with international humanitarian law, taking appropriate steps to control proliferation, and working with partners and Allies to develop standards related to the development and use of such weapon systems. The United States has actively participated in ongoing international discussion on Lethal Autonomous Weapon Systems, and anticipates continued robust international discussion of these potential weapon systems. Agencies across the U.S. Government are working to develop a single, government-wide policy, consistent with international humanitarian law, on autonomous and semi-autonomous weapons.

**Preparing for the Future**

AI holds the potential to be a major driver of economic growth and social progress, if industry, civil society, government, and the public work together to support development of the technology with thoughtful attention to its potential and to managing its risks.

The U.S. Government has several roles to play. It can convene conversations about important issues and help to set the agenda for public debate. It can monitor the safety and fairness of applications as they develop, and adapt regulatory frameworks to encourage innovation while protecting the public. It can provide public policy tools to ensure that disruption in the means and methods of work enabled by AI increases productivity while avoiding negative economic consequences for certain sectors of the workforce. It can support basic research and the application of AI to public good. It can support development of a skilled, diverse workforce. And government can use AI itself to serve the public faster,
more effectively, and at lower cost. Many areas of public policy, from education and the economic safety net, to defense, environmental preservation, and criminal justice, will see new opportunities and new challenges driven by the continued progress of AI. The U.S. Government must continue to build its capacity to understand and adapt to these changes.

As the technology of AI continues to develop, practitioners must ensure that AI-enabled systems are governable; that they are open, transparent, and understandable; that they can work effectively with people; and that their operation will remain consistent with human values and aspirations. Researchers and practitioners have increased their attention to these challenges, and should continue to focus on them.

Developing and studying machine intelligence can help us better understand and appreciate our human intelligence. Used thoughtfully, AI can augment our intelligence, helping us chart a better and wiser path forward.

A full list of the recommendations made in this report is on page 40.
Introduction

Artificial Intelligence (AI) has the potential to help address some of the biggest challenges that society faces. Smart vehicles may save hundreds of thousands of lives every year worldwide, and increase mobility for the elderly and those with disabilities. Smart buildings may save energy and reduce carbon emissions. Precision medicine may extend life and increase quality of life. Smarter government may serve citizens more quickly and precisely, better protect those at risk, and save money. AI-enhanced education may help teachers give every child an education that opens doors to a secure and fulfilling life. These are just a few of the potential benefits if the technology is developed with an eye to its benefits and with careful consideration of its risks and challenges.

The United States has been at the forefront of foundational research in AI, primarily supported for most of the field’s history by Federal research funding and work at government laboratories. The Federal Government’s support for unclassified AI R&D is managed through the Networking and Information Technology Research and Development (NITRD) program, and supported primarily by the Defense Advanced Research Projects Agency (DARPA), the National Science Foundation (NSF), the National Institutes of Health (NIH), the Office of Naval Research (ONR), and the Intelligence Advanced Research Projects Activity (IARPA). Major national research efforts such as the National Strategic Computing Initiative, the Big Data Initiative, and the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative also contribute indirectly to the progress of AI research. The current and projected benefits of AI technology are large, adding to the Nation’s economic vitality and to the productivity and well-being of its people. A companion document lays out a strategic plan for Federally-funded research and development in AI.

As a contribution toward preparing the United States for a future in which AI plays a growing role, we survey the current state of AI, its existing and potential applications, and the questions that progress in AI raise for society and public policy. We also make recommendations for specific further actions by Federal agencies and other actors.

A Brief History of AI

Endowing computers with human-like intelligence has been a dream of computer experts since the dawn of electronic computing. Although the term “Artificial Intelligence” was not coined until 1956, the roots of the field go back to at least the 1940s, and the idea of AI was crystalized in Alan Turing’s famous 1950 paper, “Computing Machinery and Intelligence.” Turing’s paper posed the question: “Can machines think?” It also proposed a test for answering that question, and raised the possibility that a machine might be programmed to learn from experience much as a young child does.

In the ensuing decades, the field of AI went through ups and downs as some AI research problems proved more difficult than anticipated and others proved insurmountable with the technologies of the time. It wasn’t until the late 1990s that research progress in AI began to accelerate, as researchers focused more on sub-problems of AI and the application of AI to real-world problems such as image recognition and medical diagnosis. An early milestone was the 1997 victory of IBM’s chess-playing computer Deep Blue

---


4 Restated in modern terms, the “Turing Test” puts a human judge in a text-based chat room with either another person or a computer. The human judge can interrogate the other party and carry on a conversation, and then the judge is asked to guess whether the other party is a person or a computer. If a computer can consistently fool human judges in this game, then the computer is deemed to be exhibiting intelligence.
over world champion Garry Kasparov. Other significant breakthroughs included DARPA’s Cognitive Agent that Learns and Organizes (CALO), which led to Apple Inc.’s Siri; IBM’s question-answering computer Watson’s victory in the TV game show “Jeopardy!”; and the surprising success of self-driving cars in the DARPA Grand Challenge competitions in the 2000s.

The current wave of progress and enthusiasm for AI began around 2010, driven by three factors that built upon each other: the availability of big data from sources including e-commerce, businesses, social media, science, and government; which provided raw material for dramatically improved machine learning approaches and algorithms; which in turn relied on the capabilities of more powerful computers. During this period, the pace of improvement surprised AI experts. For example, on a popular image recognition challenge that has a 5 percent human error rate according to one error measure, the best AI result improved from a 26 percent error rate in 2011 to 3.5 percent in 2015.

Simultaneously, industry has been increasing its investment in AI. In 2016, Google Chief Executive Officer (CEO) Sundar Pichai said, “Machine learning [a subfield of AI] is a core, transformative way by which we’re rethinking how we’re doing everything. We are thoughtfully applying it across all our products, be it search, ads, YouTube, or Play. And we’re in early days, but you will see us—in a systematic way—apply machine learning in all these areas.” This view of AI broadly impacting how software is created and delivered was widely shared by CEOs in the technology industry, including Ginni Rometty of IBM, who has said that her organization is betting the company on AI.

What is Artificial Intelligence?

There is no single definition of AI that is universally accepted by practitioners. Some define AI loosely as a computerized system that exhibits behavior that is commonly thought of as requiring intelligence. Others define AI as a system capable of rationally solving complex problems or taking appropriate actions to achieve its goals in whatever real world circumstances it encounters.

Experts offer differing taxonomies of AI problems and solutions. A popular AI textbook used the following taxonomy: (1) systems that think like humans (e.g., cognitive architectures and neural networks); (2) systems that act like humans (e.g., pass the Turing test via natural language processing; knowledge representation, automated reasoning, and learning), (3) systems that think rationally (e.g.,

---


6 The ImageNet Large Scale Visual Recognition Challenge provides a set of photographic images and asks for an accurate description of what is depicted in each image. Statistics in the text refer to the “classification error” metric in the “classiﬁcation+localization with provided training data” task. See http://image-net.org/challenges/LSVRC/.


8 See, e.g., Andrew Nusca, “IBM’s CEO Thinks Every Digital Business Will Become a Cognitive Computing Business,” Fortune, June 1 2016. ("[IBM] CEO Ginni Rometty is optimistic that the company’s wager on ‘cognitive computing,’ the term it uses for applied artificial intelligence and machine learning technologies, is the biggest bet the company will make in its 105-year history.")

logic solvers, inference, and optimization); and (4) systems that act rationally (e.g., intelligent software agents and embodied robots that achieve goals via perception, planning, reasoning, learning, communicating, decision-making, and acting). Separately, venture capitalist Frank Chen broke down the problem space of AI into five general categories: logical reasoning, knowledge representation, planning and navigation, natural language processing, and perception. And AI researcher Pedro Domingos ascribed AI researchers to five “tribes” based on the methods they use: “symbolists” use logical reasoning based on abstract symbols; “connectionists” build structures inspired by the human brain; “evolutionaries” use methods inspired by Darwinian evolution; “Bayesians” use probabilistic inference; and “analogizers” extrapolate from similar cases seen previously.

This diversity of AI problems and solutions, and the foundation of AI in human evaluation of the performance and accuracy of algorithms, makes it difficult to clearly define a bright-line distinction between what constitutes AI and what does not. For example, many techniques used to analyze large volumes of data were developed by AI researchers and are now identified as “Big Data” algorithms and systems. In some cases, opinion may shift, meaning that a problem is considered as requiring AI before it has been solved, but once a solution is well known it is considered routine data processing. Although the boundaries of AI can be uncertain and have tended to shift over time, what is important is that a core objective of AI research and applications over the years has been to automate or replicate intelligent behavior.

**The Current State of AI**

Remarkable progress has been made on what is known as Narrow AI, which addresses specific application areas such as playing strategic games, language translation, self-driving vehicles, and image recognition. Narrow AI underpins many commercial services such as trip planning, shopper recommendation systems, and ad targeting, and is finding important applications in medical diagnosis, education, and scientific research. These have all had significant societal benefits and have contributed to the economic vitality of the Nation.

General AI (sometimes called Artificial General Intelligence, or AGI) refers to a notional future AI system that exhibits apparently intelligent behavior at least as advanced as a person across the full range of cognitive tasks. A broad chasm seems to separate today’s Narrow AI from the much more difficult challenge of General AI. Attempts to reach General AI by expanding Narrow AI solutions have made little headway over many decades of research. The current consensus of the private-sector expert community, with which the NSTC Committee on Technology concurs, is that General AI will not be achieved for at least decades.

---


12 Narrow AI is not a single technical approach, but rather a set of discrete problems whose solutions rely on a toolkit of AI methods along with some problem-specific algorithms. The diversity of Narrow AI problems and solutions, and the apparent need to develop specific methods for each Narrow AI application, has made it infeasible to “generalize” a single Narrow AI solution to produce intelligent behavior of general applicability.


14 Expert opinion on the expected arrival date of AGI ranges from 2030 to centuries from now. There is a long history of excessive optimism about AI. For example, AI pioneer Herb Simon predicted in 1957 that computers
People have long speculated on the implications of computers becoming more intelligent than humans. Some predict that a sufficiently intelligent AI could be tasked with developing even better, more intelligent systems, and that these in turn could be used to create systems with yet greater intelligence, and so on, leading in principle to an “intelligence explosion” or “singularity” in which machines quickly race far ahead of humans in intelligence.\(^\text{15}\) In a dystopian vision of this process, these super-intelligent machines would exceed the ability of humanity to understand or control. If computers could exert control over many critical systems, the result could be havoc, with humans no longer in control of their destiny at best and extinct at worst. This scenario has long been the subject of science fiction stories, and recent pronouncements from some influential industry leaders have highlighted these fears.

A more positive view of the future held by many researchers sees instead the development of intelligent systems that work well as helpers, assistants, trainers, and teammates of humans, and are designed to operate safely and ethically.

The NSTC Committee on Technology’s assessment is that long-term concerns about super-intelligent General AI should have little impact on current policy. The policies the Federal Government should adopt in the near-to-medium term if these fears are justified are almost exactly the same policies the Federal Government should adopt if they are not justified. The best way to build capacity for addressing the longer-term speculative risks is to attack the less extreme risks already seen today, such as current security, privacy, and safety risks, while investing in research on longer-term capabilities and how their challenges might be managed. Additionally, as research and applications in the field continue to mature, practitioners of AI in government and business should approach advances with appropriate consideration of the long-term societal and ethical questions – in additional to just the technical questions – that such advances portend. Although prudence dictates some attention to the possibility that harmful super-intelligence might someday become possible, these concerns should not be the main driver of public policy for AI.

**Machine Learning**

Machine learning is one of the most important technical approaches to AI and the basis of many recent advances and commercial applications of AI. Modern machine learning is a statistical process that starts with a body of data and tries to derive a rule or procedure that explains the data or can predict future data. This approach—learning from data—contrasts with the older “expert system” approach to AI, in which programmers sit down with human domain experts to learn the rules and criteria used to make decisions, and translate those rules into software code. An expert system aims to emulate the principles used by human experts, whereas machine learning relies on statistical methods to find a decision procedure that works well in practice.

An advantage of machine learning is that it can be used even in cases where it is infeasible or difficult to write down explicit rules to solve a problem. For example, a company that runs an online service might use machine learning to detect user log-in attempts that are fraudulent. The company might start with a large data set of past login attempts, with each attempt labeled as fraudulent or not using the benefit of

---

\(^\text{15}\) It is far from certain that this sort of explosive growth in intelligence is likely, or even possible. Another plausible extrapolation from current knowledge is that machine intelligence will continue to increase gradually even after surpassing human intelligence.
hindsight. Based on this data set, the company could use machine learning to derive a rule to apply to future login attempts that predicts which attempts are more likely to be fraudulent and should be subjected to extra security measures. In a sense, machine learning is not an algorithm for solving a specific problem, but rather a more general approach to finding solutions for many different problems, given data about them.

To apply machine learning, a practitioner starts with a historical data set, which the practitioner divides into a *training set* and a *test set*. The practitioner chooses a *model*, or mathematical structure that characterizes a range of possible decision-making rules with adjustable *parameters*. A common analogy is that the model is a “box” that applies a rule, and the parameters are adjustable knobs on the front of the box that control how the box operates. In practice, a model might have many millions of parameters.

The practitioner also defines an *objective function* used to evaluate the desirability of the outcome that results from a particular choice of parameters. The objective function will typically contain parts that reward the model for closely matching the training set, as well as parts that reward the use of simpler rules.

*Training* the model is the process of adjusting the parameters to maximize the objective function. Training is the difficult technical step in machine learning. A model with millions of parameters will have astronomically more possible outcomes than any algorithm could ever hope to try, so successful training algorithms have to be clever in how they explore the space of parameter settings so as to find very good settings with a feasible level of computational effort.

Once a model has been trained, the practitioner can use the test set to evaluate the accuracy and effectiveness of the model. The goal of machine learning is to create a trained model that will *generalize*—it will be accurate not only on examples in the training set, but also on future cases that it has never seen before. While many of these models can achieve better-than-human performance on narrow tasks such as image labeling, even the best models can fail in unpredictable ways. For example, for many image labeling models it is possible to create images that clearly appear to be random noise to a human but will be falsely labeled as a specific object with high confidence by a trained model.

Another challenge in using machine learning is that it is typically not possible to extract or generate a straightforward explanation for why a particular trained model is effective. Because trained models have a very large number of adjustable parameters—often hundreds of millions or more—training may yield a model that "works," in the sense of matching the data, but is not necessarily the simplest model that works. In human decision-making, any opacity in the process is typically due to not having enough information about why a decision was reached, because the decider may be unable to articulate why the decision “felt right.” With machine learning, everything about the decision procedure is known with mathematical precision, but there may be simply too much information to interpret clearly.

**Deep Learning**

In recent years, some of the most impressive advancements in machine learning have been in the subfield of deep learning, also known as deep network learning. Deep learning uses structures loosely inspired by the human brain, consisting of a set of units (or “neurons”). Each unit combines a set of input values to produce an output value, which in turn is passed on to other neurons downstream. For example, in an image recognition application, a first layer of units might combine the raw data of the image to recognize simple patterns in the image; a second layer of units might combine the results of the first layer to recognize patterns-of-patterns; a third layer might combine the results of the second layer; and so on.

---

Deep learning networks typically use many layers—sometimes more than 100—and often use a large number of units at each layer, to enable the recognition of extremely complex, precise patterns in data.

In recent years, new theories of how to construct and train deep networks have emerged, as have larger, faster computer systems, enabling the use of much larger deep learning networks. The dramatic success of these very large networks at many machine learning tasks has come as a surprise to some experts, and is the main cause of the current wave of enthusiasm for machine learning among AI researchers and practitioners.

**Autonomy and Automation**

AI is often applied to systems that can control physical actuators or trigger online actions. When AI comes into contact with the everyday world, issues of autonomy, automation, and human-machine teaming arise.

*Autonomy* refers to the ability of a system to operate and adapt to changing circumstances with reduced or without human control. For example, an autonomous car could drive itself to its destination. Despite the focus in much of the literature on cars and aircraft, autonomy is a much broader concept that includes scenarios such as automated financial trading and automated content curation systems. Autonomy also includes systems that can diagnose and repair faults in their own operation, such as identifying and fixing security vulnerabilities.

*Automation* occurs when a machine does work that might previously have been done by a person. The term relates to both physical work and mental or cognitive work that might be replaced by AI. Automation, and its impact on employment, have been significant social and economic phenomena since at least the Industrial Revolution. It is widely accepted that AI will automate some jobs, but there is more debate about whether this is just the next chapter in the history of automation or whether AI will affect the economy differently than past waves of automation have previously.

**Human-Machine Teaming**

In contrast to automation, where a machine substitutes for human work, in some cases a machine will complement human work. This may happen as a side-effect of AI development, or a system might be developed specifically with the goal of creating a human-machine team. Systems that aim to complement human cognitive capabilities are sometimes referred to as intelligence augmentation.

In many applications, a human-machine team can be more effective than either one alone, using the strengths of one to compensate for the weaknesses of the other. One example is in chess playing, where a weaker computer can often beat a stronger computer player, if the weaker computer is given a human teammate—this is true even though top computers are much stronger players than any human. Another example is in radiology. In one recent study, given images of lymph node cells, and asked to determine

---

17 Different definitions of “automation” are used in different settings. The definition used in the main text, involving the substitution of machine labor for human labor, is commonly used in economics. Another definition is used in the systems analysis setting in the Department of Defense (DoD): Automation means that the system functions with little or no human operator involvement. However the system performance is limited to the specific pre-programmed actions it has been designed to execute. Once the system is initiated by a human operator, it executes its task according to those instructions and subroutines, which have been tested and validated. Typically these are well-defined tasks that have predetermined responses, i.e., rule-based responses in reasonably well-known and structured environments.

whether or not the cells contained cancer, an AI-based approach had a 7.5 percent error rate, where a human pathologist had a 3.5 percent error rate; a combined approach, using both AI and human input, lowered the error rate to 0.5 percent, representing an 85 percent reduction in error.\footnote{Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, Andrew H. Beck, “Deep Learning for Identifying Metastatic Breast Cancer,” June 18, 2016, https://arxiv.org/pdf/1606.05718v1.pdf.}
Public Outreach and Development of this Report

This report was developed by the NSTC’s Subcommittee on Machine Learning and Artificial Intelligence, which was chartered in May 2016 to foster interagency coordination and provide technical and policy advice on topics related to AI, and to monitor the development of AI technologies across industry, the research community, and the Federal Government. The report follows a series of public outreach activities led by OSTP, designed to allow government officials thinking about these topics to learn from experts and from the public. This public outreach on AI included five co-hosted public workshops, and a public Request for Information (RFI). The public workshops were:

- AI, Law, and Governance (May 24, in Seattle, co-hosted by OSTP, the National Economic Council (NEC), and the University of Washington);
- AI for Social Good (June 7, in Washington DC, co-hosted by OSTP, the Association for the Advancement of AI (AAAI) and the Computing Community Consortium (CCC));
- Future of AI: Emerging Topics and Societal Benefit at the Global Entrepreneurship Summit (June 23, in Palo Alto, co-hosted by OSTP and Stanford University);
- AI Technology, Safety, and Control (June 28, in Pittsburgh, co-hosted by OSTP and Carnegie Mellon University); and
- Social and Economic Impacts of AI (July 7, in New York, co-hosted by OSTP, NEC, and New York University).

In conjunction with each of the five workshops, the private-sector co-hosts organized separate meetings or conference sessions which government staff attended. Total in-person attendance at the public events was more than 2,000 people, in addition to international online streaming audiences, which included more than 3,500 people for the Washington, DC workshop livestream alone.

OSTP also published a Request for Information (RFI) seeking public comment on the topics of the workshops. The RFI closed on July 22, 2016 and received 161 responses. Comments submitted in response to the public RFI were published by OSTP on September 6, 2016.20

---

Applications of AI for Public Good

One area of great optimism about AI and machine learning is their potential to improve people’s lives by helping to solve some of the world’s greatest challenges and inefficiencies. The promise of AI has been compared to the transformative impacts of advances in mobile computing. Public- and private-sector investments in basic and applied R&D on AI have already begun reaping major benefits for the public in fields as diverse as health care, transportation, the environment, criminal justice, and economic inclusion.

At Walter Reed Medical Center, the Department of Veteran Affairs is using AI to better predict medical complications and improve treatment of severe combat wounds, leading to better patient outcomes, faster healing, and lower costs. The same general approach—predicting complications to enable preventive treatment—has also reduced hospital-acquired infections at Johns Hopkins University. Given the current transition to electronic health records, predictive analysis of health data may play a key role across many health domains like precision medicine and cancer research.

In transportation, AI-enabled smarter traffic management applications are reducing wait times, energy use, and emissions by as much as 25 percent in some places. Cities are now beginning to leverage the type of responsive dispatching and routing used by ride-hailing services, and linking it with scheduling and tracking software for public transportation to provide just-in-time access to public transportation that can often be faster, cheaper and, in many cases, more accessible to the public.

Some researchers are leveraging AI to improve animal migration tracking by using AI image classification software to analyze tourist photos from public social media sites. The software can identify individual animals in the photos and build a database of their migration using the data and location stamps on the photos. At OSTP’s AI for Social Good workshop, researchers talked about building some of the largest available datasets to date on the populations and migrations of whales and large African animals, and about launching a project to track “The Internet of Turtles” to gain new insights about sea life. Other speakers described uses of AI to optimize the patrol strategy of anti-poaching agents, and to design habitat preservation strategies to maximize the genetic diversity of endangered populations.

---


Autonomous sailboats and watercraft are already patrolling the oceans carrying sophisticated sensor instruments, collecting data on changes in Arctic ice and sensitive ocean ecosystems in operations that would be too expensive or dangerous for crewed vessels. Autonomous watercraft may be much cheaper to operate than manned ships, and may someday be used for enhanced weather prediction, climate monitoring, or policing illegal fishing.27

AI also has the potential to improve aspects of the criminal justice system, including crime reporting, policing, bail, sentencing, and parole decisions. The Administration is exploring how AI can responsibly benefit current initiatives such as Data Driven Justice and the Police Data Initiative that seek to provide law enforcement and the public with data that can better inform decision-making in the criminal justice system, while also taking care to minimize the possibility that AI might introduce bias or inaccuracies due to deficiencies in the available data.

Several U.S. academic institutions have launched initiatives to use AI to tackle economic and social challenges. For example, the University of Chicago created an academic program that uses data science and AI to address public challenges such as unemployment and school dropouts.28 The University of Southern California launched the Center for Artificial Intelligence in Society, an institute dedicated to studying how computational game theory, machine learning, automated planning and multi-agent reasoning techniques can help to solve socially relevant problems like homelessness. Meanwhile, researchers at Stanford University are using machine learning in efforts to address global poverty by using AI to analyze satellite images of likely poverty zones to identify where help is needed most.29

Many uses of AI for public good rely on the availability of data that can be used to train machine learning models and test the performance of AI systems. Agencies and organizations with data that can be released without implicating personal privacy or trade secrets can help to enable the development of AI by making those data available to researchers. Standardizing data schemas and formats can reduce the cost and difficulty of making new data sets useful.

Recommendation 1: Private and public institutions are encouraged to examine whether and how they can responsibly leverage AI and machine learning in ways that will benefit society. Social justice and public policy institutions that do not typically engage with advanced technologies and data science in their work should consider partnerships with AI researchers and practitioners that can help apply AI tactics to the broad social problems these institutions already address in other ways.

Recommendation 2: Federal agencies should prioritize open training data and open data standards in AI. The government should emphasize the release of datasets that enable the use of AI to address social challenges. Potential steps may include developing an “Open Data for AI” initiative with the objective of releasing a significant number of government data sets to accelerate AI research and galvanize the use of open data standards and best practices across government, academia, and the private sector.


AI in the Federal Government

The Administration is working to develop policies and internal practices that will maximize the economic and societal benefits of AI and promote innovation. These policies and practices may include:

- investing in basic and applied research and development (R&D);
- serving as an early customer for AI technologies and their applications;
- supporting pilot projects and creating testbeds in real-world settings;
- making data sets available to the public;
- sponsoring incentive prizes;
- identifying and pursuing Grand Challenges to set ambitious but achievable goals for AI;
- funding rigorous evaluations of AI applications to measure their impact and cost-effectiveness; and
- creating a policy, legal, and regulatory environment that allows innovation to flourish while protecting the public from harm.

Using AI in Government to Improve Services and Benefit the American People

One challenge in using AI to improve services is that the Federal Government’s capacity to foster and harness innovation in order to better serve the country varies widely across agencies. Some agencies are more focused on innovation, particularly those agencies with large R&D budgets, a workforce that includes many scientists and engineers, a culture of innovation and experimentation, and strong ongoing collaborations with private-sector innovators. Many also have organizations that are specifically tasked with supporting high-risk, high-return research (e.g., the advanced research projects agencies in the Departments of Defense and Energy, as well as the Intelligence Community), and fund R&D across the full range from basic research to advanced development. Other agencies like the NSF have research and development as their primary mission.

But some agencies, particularly those charged with reducing poverty and increasing economic and social mobility, have more modest levels of relevant capabilities, resources, and expertise.\(^{30}\) For example, while the National Institutes of Health (NIH) has an R&D budget of more than $30 billion, the Department of Labor’s R&D budget is only $14 million. This limits the Department of Labor’s capacity to explore applications of AI, such as applying AI-based “digital tutor” technology to increase the skills and incomes of non-college educated workers.

DARPA’s “Education Dominance” program serves as an example of AI’s potential to fulfill and accelerate agency priorities. DARPA, intending to reduce from years to months the time required for new Navy recruits to become experts in technical skills, now sponsors the development of a digital tutor that uses AI to model the interaction between an expert and a novice. An evaluation of the digital tutor program concluded that Navy recruits using the digital tutor to become IT systems administrators frequently outperform Navy experts with 7-10 years of experience in both written tests of knowledge and real-world problem solving.\(^{31}\)

Preliminary evidence based on digital tutor pilot projects also suggests that workers who have completed a training program that uses the digital tutor are more likely to get a high-tech job that dramatically

---


increases their incomes. These wage increases appear to be much larger than the impacts of current workforce development programs. Ideally, these results would be confirmed with independently conducted, randomized, controlled trials. Currently, the cost of developing digital tutors is high, and there is no repeatable methodology for developing effective digital tutors. Research that enables the emergence of an industry that uses AI approaches such as digital tutors could potentially help workers acquire in-demand skills.

Recommendation 3: The Federal Government should explore ways to improve the capacity of key agencies to apply AI to their missions. For example, Federal agencies should explore the potential to create DARPA-like organizations to support high-risk, high-reward AI research and its application, much as the Department of Education has done through its proposal to create an “ARPA-ED,” to support R&D to determine whether AI and other technologies could significantly improve student learning outcomes.

Recommendation 4: The NSTC MLAI subcommittee should develop a community of practice for AI practitioners across government. Agencies should work together to develop and share standards and best practices around the use of AI in government operations. Agencies should ensure that Federal employee training programs include relevant AI opportunities.

---


AI and Regulation

AI has applications in many products, such as cars and aircraft, which are subject to regulation designed to protect the public from harm and ensure fairness in economic competition. How will the incorporation of AI into these products affect the relevant regulatory approaches? In general, the approach to regulation of AI-enabled products to protect public safety should be informed by assessment of the aspects of risk that the addition of AI may reduce, alongside the aspects of risk that it may increase. If a risk falls within the bounds of an existing regulatory regime, moreover, the policy discussion should start by considering whether the existing regulations already adequately address the risk, or whether they need to be adapted to the addition of AI. Also, where regulatory responses to the addition of AI threaten to increase the cost of compliance or slow the development or adoption of beneficial innovations, policymakers should consider how those responses could be adjusted to lower costs and barriers to innovation without adversely impacting safety or market fairness.

The general consensus of the RFI commenters was that broad regulation of AI research or practice would be inadvisable at this time. Instead, commenters said that the goals and structure of existing regulations were sufficient, and commenters called for existing regulation to be adapted as necessary to account for the effects of AI. For example, commenters suggested that motor vehicle regulation should evolve to account for the anticipated arrival of autonomous vehicles, and that the necessary evolution could be carried out within the current structure of vehicle safety regulation. In doing so, agencies must remain mindful of the fundamental purposes and goals of regulation to safeguard the public good, while creating space for innovation and growth in AI.

Effective regulation of technologies such as AI requires agencies to have in-house technical expertise to help guide regulatory decision-making. The need for senior-level expert participation exists at regulating departments and agencies, and at all stages of the regulatory process. A range of personnel assignment and exchange models (e.g. hiring authorities) can be used to develop a Federal workforce with more diverse perspectives on the current state of technological development. One example of such an authority is the Intergovernmental Personnel Act (IPA) Mobility Program, which provides for the temporary assignment of personnel between the Federal Government and state and local governments, colleges and universities, Indian tribal governments, federally funded research and development centers, and other eligible organizations. If used strategically, the IPA program can help agencies meet their needs for hard-to-fill positions and increase their ability to hire candidates from diverse technical backgrounds. Federal employees serving in IPA assignments can serve as both recruiters and ambassadors for the Federal workforce. For example, agency staff sent to colleges and universities as instructors can inspire students to consider Federal employment. Likewise, programs that rotate employees through different jobs and sectors can help government employees gain knowledge and experience to inform regulation and policy, especially as it relates to emergent technologies like AI.

34 Ed Felten and Terah Lyons, “Public Input and Next Steps on the Future of Artificial Intelligence.”
Case Study: Autonomous Vehicles and Aircraft

A relevant example of the regulatory challenges associated with an agency updating legacy regulations to account for new AI-based products is the work of the Department of Transportation (DOT) on automated vehicles and unmanned aircraft systems (UAS, or “drones”). Within DOT, automated cars are regulated by the National Highway Traffic Safety Administration (NHTSA) and aircraft are regulated by the Federal Aviation Administration (FAA).

The Promise of Autonomy

The application of AI to vehicles and aircraft has captured the public imagination. Today’s new cars have AI-based driver assist features like self-parking and advanced cruise controls that keep a car in its lane and adjust speed based on surrounding vehicles. Experimental fully automated cars monitored by humans can already be seen driving on the roads. The consensus of experts is that automated surface vehicle technology will eventually be safer than human drivers and may someday prevent most of the tens of thousands of fatalities that occur annually on the Nation’s roads and highways.

Automated vehicles also offer the possibility of greater mobility for the elderly and Americans with disabilities who may not be able to drive. First- and last-mile access to transit and other novel transportation approaches may provide communities isolated from essential services such as jobs, health care, and groceries unprecedented access to opportunity. A well-designed system of automated vehicles able to predict and avoid collisions may also significantly reduce transportation-related emissions and energy consumption. The Administration is taking steps to make this vision a reality, including the proposed $3.9 billion investment in the President’s Fiscal Year (FY) 2017 Budget by the Department of Transportation in automated and connected vehicle research, development, and deployment efforts, to ensure that the United States maintains its lead in automated vehicle technologies.35

Moving to the air, since the early 1990s, commercial UAS have operated on a limited basis in the National Airspace System (NAS). Until recently, UAS mainly supported government operations, such as military and border security operations. But in recent years, potential applications have rapidly expanded to include aerial photography, surveying land and crops, monitoring forest fires, responding to disasters, and inspecting critical infrastructure. Several government agencies are already operating UAS to advance their missions, and thousands of Americans have obtained the necessary authority from the Federal Aviation Administration (FAA) for commercial UAS operations, a process accelerated under the FAA’s “Small UAS Rule” that took effect in August 2016 and the FAA’s Small UAS Aircraft Registration Service that launched in December 2015. The FAA estimates that the number of UAS registered for commercial use will exceed 600,000 by August 2017.

One estimate of the economic impact of integrating of UAS into the airspace predicted more than $13.6 billion of economic value created by UAS in the first three years of integration, with sustainable growth predicted to follow. A 2013 study from the Association for Unmanned Vehicle Systems International predicted that the commercial drone industry could generate more than $82 billion for the U.S. economy and create more than 100,000 new jobs over the next 10 years. Tax revenue to the states was predicted to increase by more than $482 million in the first decade after integration.

### Ensuring Safety

Realizing the potential benefits of these promising technologies requires that government take steps to ensure the safety of the airspace and roads, while continuing to foster a culture of innovation and growth. The United States has the safest and most complex aviation system in the world, and the public relies on FAA oversight to establish safety standards. Federal Motor Vehicle Safety Standards (FMVSS) place requirements on manufacturers to develop safe surface vehicles, and NHTSA has the authority to recall vehicles in the event of an unreasonable risk to safety. While there is considerable opportunity to reduce the fatalities on roads and highways, current practices result in approximately one fatality for every 100 million vehicle miles traveled. Equaling or exceeding such performance in automated vehicles is a formidable challenge.

Applying techniques of AI in such safety-critical environments raises several challenges. First among these is the need to translate human responsibilities while driving or flying into software. Unlike in some other successful applications of Narrow AI, there are no concise descriptions for the task of operating ground or air vehicles. Each of these operations is multifaceted, with responsibilities including guiding the vehicle, detecting and avoiding obstacles, and handing mechanical failures such as flat tires. While subtasks such as navigation or certain types of perception may align with certain existing Narrow AI solutions, the integration and prioritization of these tasks may not. It may seem straightforward to simply obey all traffic laws, but a skilled human driver may cross a double-yellow road boundary to avoid an accident or move past a double-parked vehicle. Though such situations may be rare, they cannot be

---

36 The National Airspace System is the network of air navigation facilities, air traffic control facilities, airports, technology, and rules and regulations that are needed to protect persons and property on the ground, and to establish a safe and efficient airspace environment for civil, commercial, and military aviation.


39 Ibid.
ignored—simple arithmetic dictates that in order for failures to occur at least as infrequently as they do with human drivers, a system must handle many such rare cases without failure.

For systems that rely on machine learning, the need to get rare cases right has implications for system design and testing. Machine learning approaches can be more confident that a case will be handled correctly if a similar case is in the training set. The challenge is how to develop a data set that includes enough of the rare cases that contribute to the risk of an accident. Commercial aviation has mechanisms for sharing incident and safety data across the industry, but reporting may not be second nature to recently credentialed UAS operators who are new to the safety and accountability culture of the traditional aviation industry. No comparable system currently exists for the automotive industry—only fatal accidents are reported, and the collection and reporting of other traffic safety information is done, if at all, in a disparate manner at the state or local level. The lack of consistently reported incident or near-miss data increases the number of miles or hours of operation necessary to establish system safety, presenting an obstacle to certain AI approaches that require extensive testing for validation.

To facilitate safe testing, the FAA has designated six UAS Test Sites across the country and provided blanket authorization for UAS operations within these sites. Activities at the sites include a project to extend NASA’s multi-year research on UAS traffic management (UTM) to identify operational requirements for large-scale beyond visual line-of-sight UAS operations in low-altitude airspace. Similarly, ground vehicle testbeds such as the Connected Vehicle Pilots and the deployment of automated vehicles in Columbus, Ohio, winner of the Department of Transportation’s $40 million Smart City Challenge in 2016, will provide rich baseline and interaction data for AI researchers.

**Recommendation 7:** The Department of Transportation should work with industry and researchers on ways to increase sharing of data for safety, research, and other purposes. In light of the future importance of AI in surface and air, Federal actors should focus in the near-term on developing increasingly rich sets of data, consistent with consumer privacy, that can better inform policy-making as these technologies mature.

### Adapting Current Regulations

While the regulatory approaches for the Nation’s airspace and highways differ, the approaches to integrating autonomous vehicles and aircraft share a common goal: both the FAA and NHTSA are working to establish nimble and flexible frameworks that ensure safety while encouraging innovation.

With respect to airspace regulation, a significant step toward enabling the safe integration of UAS into the airspace was the FAA’s promulgation of the Part 107, or “Small UAS,” final rule, which took effect on August 29, 2016. For the first time, the rule authorizes widespread non-recreational flights of UAS under 55 pounds. The rule limits flights to daytime, at an altitude of 400 feet or less, with the vehicle controlled by a licensed operator and within the operator’s direct line of sight. Flights over people are not allowed. Subsequent rules are planned, to relax these restrictions as experience and data show how to do so safely. In particular, DOT is currently developing a Notice of Proposed Rulemaking proposing a regime for certain types of “micro UAS” to conduct operations over people, with a rule on expanded operations expected to follow.

The FAA has not yet publicly announced a clear path to a regulation allowing fully autonomous flight. Though safe integration of autonomous aircraft into the airspace will be a complex process, the FAA is

---

40 This report uses the term “autonomous” for an aircraft that is controlled by a machine rather than a human. “Piloted” refers to an aircraft that has a human onboard who is controlling the aircraft. “Remotely-piloted” refers to
preparing for a not-so-distant technological future in which autonomous and piloted aircraft fly together in a seamlessly integrated airspace system.

New approaches to airspace management may also include AI-based enhancement of the air traffic control system. Projected future air traffic densities and diversity of operations are unlikely to be feasible within the current airspace management architecture, due to current limits on air/ground integration, and reliance on human-to-human communication in air and ground practices. The cost of U.S. air transportation delays in 2007, the latest year for which there is reliable public data, was estimated to be $31.2 billion—a number that has presumably grown as user volume has increased since that year. Though some flight delays are unavoidable due to weather and other constraints, adopting new aviation technologies, enabling policies, and infrastructure upgrades could significantly increase efficiency of operation in the U.S. airspace. Such solutions include AI and machine learning-based architectures that have the potential to better accommodate a wider range of airspace users, including piloted and unpiloted aircraft, and to use airspace more efficiently without undermining safety. Development and deployment of such technologies would help ensure global competitiveness for airspace users and service providers, while increasing safety and reducing cost.

With respect to surface transportation, the most significant step currently underway to establish a common framework is the Federal Automated Vehicles Policy that the Administration released on September 20, 2016. The policy had several parts:

- guidance for manufacturers, developers, and other organizations outlining a 15 point “Safety Assessment” for the safe design, development, testing, and deployment of highly automated vehicles;
- a model state policy, which clearly distinguishes Federal and State responsibilities and recommends policy areas for states to consider, with a goal of generating a consistent national framework for the testing and operation of automated vehicles, while leaving room for experimentation by states;
- an analysis of current regulatory tools that NHTSA can use to aid the safe development of automated vehicles, such as interpreting current rules to allow for appropriate flexibility in design, providing limited exemptions to allow for testing of nontraditional vehicle designs, and ensuring that unsafe automated vehicles are removed from the road; and

an aircraft that is controlled by a human who is not on board. “Manned” means there is a human onboard who may or may not be in control.


a discussion of new tools and authorities that the agency could consider seeking in the future to aid the safe and efficient deployment of new lifesaving technologies and ensure that technologies deployed on the road are safe.

DOT intends for the guidance and the model state policy to be routinely updated as new data are learned and research completed.

Recommendation 8: The U.S. Government should invest in developing and implementing an advanced and automated air traffic management system that is highly scalable, and can fully accommodate autonomous and piloted aircraft alike.

Recommendation 9: The Department of Transportation should continue to develop an evolving framework for regulation to enable the safe integration of fully automated vehicles and UAS, including novel vehicle designs, into the transportation system.
Research and Workforce

Government also has an important role to play in advancing the AI field by investing in research and development, developing a workforce that is skilled and diverse, and managing the economic impacts of these technologies as they develop. A separate National Artificial Intelligence Research and Development Strategic Plan is being published in conjunction with this report. This section discusses additional policy issues related to research and workforce development.

Monitoring Progress in AI

Given the potential impacts of AI, society would benefit from accurate and timely methods for monitoring and forecasting AI developments. Several projects have attempted to forecast AI futures. The 2009 AAAI Presidential Panel on Long-Term AI Futures and the 2015 Future of AI Conference brought together AI experts to predict the future of their field. In addition, Stanford’s One-Hundred Year Study on Artificial Intelligence plans to conduct “a series of periodic studies on how AI will affect automation, national security, psychology, ethics, law, privacy, democracy, and other issues.” The first of these studies was published in September 2016.

One potentially useful line of research is to survey expert judgments over time. As one example, a survey of AI researchers found that 80 percent of respondents believed that human-level General AI will eventually be achieved, and half believed it is at least 50 percent likely to be achieved by the year 2040. Most respondents also believed that General AI will eventually surpass humans in general intelligence. While these particular predictions are highly uncertain, as discussed above, such surveys of expert judgment are useful, especially when they are repeated frequently enough to measure changes in judgment over time. One way to elicit frequent judgments is to run “forecasting tournaments” such as prediction markets, in which participants have financial incentives to make accurate predictions.

45 The track record of technology forecasts, in general, and AI forecasts, in particular, suggests this may be difficult. One of the largest retrospective reviews of technology forecasts over the last 50 years found that forecasts with time horizons beyond 10 years were rarely better than coin-flips. (Carrie Mullins, “Retrospective Analysis of Technology Forecasting,” The Tauri Group, August 13, 2012.) One review of 95 timeline predictions for AI from 1950 to 2012 found that most forecasts predicted General AI would be achieved “in the next 20 years.” (Stuart Armstrong, Kaj Sotala, Seán S. ÓhÉigeartaigh, “The errors, insights and lessons of famous AI predictions – and what they mean for the future,” Journal of Experimental & Theoretical Artificial Intelligence, May 20, 2014.)


49 Ibid.


research has found that technology developments can often be accurately predicted by analyzing trends in publication and patent data⁵².

At present, the majority of basic research in AI is conducted by academics and by commercial labs that regularly announce their findings and publish them in the research literature. If competition drives commercial labs towards increased secrecy, monitoring of progress may become more difficult, and public concern may increase.

One particularly valuable line of research is to identify milestones that could represent or foreshadow significant leaps in AI capabilities. When asked during the outreach workshops and meetings how government could recognize milestones of progress in the field, especially those that indicate the arrival of General AI may be approaching, researchers tended to give three distinct but related types of answers:

1. **Success at broader, less structured tasks:** In this view, the transition from present Narrow AI to an eventual General AI will occur by gradually broadening the capabilities of Narrow AI systems so that a single system can cover a wider range of less structured tasks. An example milestone in this area would be a housecleaning robot that is as capable as a person at the full range of routine housecleaning tasks.

2. **Unification of different “styles” of AI methods:** In this view, AI currently relies on a set of separate methods or approaches, each useful for different types of applications. The path to General AI would involve a progressive unification of these methods. A milestone would involve finding a single method that is able to address a larger domain of applications that previously required multiple methods.

3. **Solving specific technical challenges, such as transfer learning:** In this view, the path to General AI does not lie in progressive broadening of scope, nor in unification of existing methods, but in progress on specific technical grand challenges, opening up new ways forward. The most commonly cited challenge is transfer learning, which has the goal of creating a machine learning algorithm whose result can be broadly applied (or transferred) to a range of new applications. For example, transfer learning might allow a model to be trained to translate English to Spanish, in such a way that the resulting model could “transfer” its knowledge to similar tasks such as Chinese to French translation, or writing poetry in Russian, enabling these new tasks to be learned much more quickly.

**Recommendation 10:** The NSTC Subcommittee on Machine Learning and Artificial Intelligence should monitor developments in AI, and report regularly to senior Administration leadership about the status of AI, especially with regard to milestones. The Subcommittee should update the list of milestones as knowledge advances and the consensus of experts changes over time. The Subcommittee should consider reporting to the public on AI developments, when appropriate.

**Recommendation 11:** The Government should monitor the state of AI in other countries, especially with respect to milestones.

**Recommendation 12:** Industry should work with government to keep government updated on the general progress of AI in industry, including the likelihood of milestones being reached soon.

Federal Support for AI Research

In 2015, the U.S. Government’s investment in unclassified R&D in AI-related technologies was approximately $1.1 billion, with preliminary estimates showing growth to $1.2 billion in 2016. Throughout the workshops and public outreach on AI conducted by OSTP, government officials heard calls for greater government investment in AI research and development, from business leaders, technologists, and economists.

Leading researchers in AI were optimistic about sustaining the recent rapid progress in AI and its application to an ever wider range of applications. At the same time they emphasized that there are many deep unanswered questions, and no clear path toward General AI.

Researchers reported that enthusiasm for and investment in AI research has fluctuated over recent decades—one low period was known as the “AI winter”—and they emphasized the importance of sustained investment given the history of major computer science advances taking 15 years or more to transition from conception in the lab to industrial maturity.

A strong case can be made in favor of increased Federal funding for research in AI. Analysis by the Council of Economic Advisers (CEA) indicates that beyond AI, across all research areas, doubling or tripling research investment would be a net positive for the Nation due to the resulting increase in economic growth. Although it may not be feasible fiscally to increase funding for all research by that amount, a targeted increase in areas of high economic and strategic value may offer many benefits with much smaller budgetary impact than an across-the-board increase. AI qualifies as a high-leverage area, and research agencies report that the AI research community can absorb a significant funding increase productively, leading to faster progress on AI and a larger cadre of trained AI practitioners. In a speech delivered at an AI workshop in New York City in July 2016, CEA Chairman Jason Furman said, “We have had substantial innovation in robotics, AI, and other areas in the last decade. But we will need a much faster pace of innovation in these areas to really move the dial on productivity growth going forward,” noting that the biggest worry that he had about AI is “that we do not have enough of [it].”

To be sure, the private sector will be the main engine of progress on AI. But as it stands, there is an underinvestment in basic research—research with long time horizons conducted for the sole purpose of furthering the scientific knowledge base—in part because it is difficult for a private firm to get a return from its investment in such research in a reasonable time frame. Basic research benefits everyone, but only the firm doing the research pays the costs. The literature suggests that, as a result, current levels of R&D spending are half to one-quarter of the level of R&D investment that would produce the optimal level of economic growth.

---


54 Jason Furman, “Is This Time Different? The Opportunities and Challenges of Artificial Intelligence.”

Workforce Development and Diversity

The rapid growth of AI has dramatically increased the need for people with relevant skills to support and advance the field. The AI workforce includes AI researchers who drive fundamental advances in AI, a larger number of specialists who refine AI methods for specific applications, and a much larger number of users who operate those applications in specific settings. For researchers, AI training is inherently interdisciplinary, often requiring a strong background in computer science, statistics, mathematical logic, and information theory. For specialists, training typically requires a background in software engineering and in the application area. For users, familiarity with AI technologies is needed to apply AI technologies reliably.

The Role of Government

The AI workforce challenge is in part a science, technology, engineering, and mathematics (STEM) education challenge that remains a priority focus of the NSTC, OSTP, and other agencies. NSF and the Department of Education are working with the private sector and across government to advance education quality, flexibility, and domain impact, to address goals such as sustained economic development, increased inclusion and diversity, and improved outcome measures. The NSTC Committee on Science, Technology and Mathematics Education (CoSTEM) brings together Federal agencies supporting STEM education programs to coordinate efforts on multiple topics, including AI education.

AI knowledge and education are increasingly emphasized in Federal STEM education programs. There are several key roles for the Federal government in AI workforce development, including supporting graduate students, funding research on AI curriculum design and impact, and accrediting AI education programs.

The Role of Schools and Universities

Integrating AI, data science, and related fields throughout the Nation’s education system is essential to developing a workforce that can address national priorities. Educational institutions are establishing and growing AI programs at all levels. Universities, colleges, and even secondary schools are expanding AI and data science curricula, but more programs and teachers are needed.

There are several key roles for academic institutions:

Recommendation 13: The Federal government should prioritize basic and long-term AI research. The Nation as a whole would benefit from a steady increase in Federal and private-sector AI R&D, with a particular emphasis on basic research and long-term, high-risk research initiatives. Because basic and long-term research especially are areas where the private sector is not likely to invest, Federal investments will be important for R&D in these areas.


56 There is also a need for the development of a strong research community in fields outside of technical disciplines related to AI, to examine the impacts and implications of AI on economics, social science, health, and other areas of research.
PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE

- building and sustaining the researcher workforce, including computer scientists, statisticians, database and software programmers, curators, librarians, and archivists with specialization in data science;
- training the specialist workforce, by emphasizing AI methods within software development courses, offering applied AI courses that demonstrate the applications of AI to other domains, and incorporating AI and data science challenges posed by industry, civil society, and government into active case studies;
- ensuring that the user workforce has the necessary familiarity with AI systems to meet the needs of users and of institutions across industry, government, and academia;
- supporting training through seed grants, professional development stipends, internships, fellowships, and summer research experiences; and
- recruiting and retaining faculty, as industrial salaries grow faster than academic salaries for skilled researchers.

Community colleges, two-year colleges, and certificate programs play an important role in providing opportunities for students and professionals to acquire necessary skills for a modest investment of their time and money. These opportunities may be especially relevant to workers expanding their skills, veterans returning to the workforce, and unemployed people seeking a way to reenter the workforce.

An AI-enabled world demands a data-literate citizenry that is able to read, use, interpret, and communicate about data, and participate in policy debates about matters affected by AI. Data science education as early as primary or secondary school can help to improve nationwide data literacy, while also preparing students for more advanced data science concepts and coursework after high school.

AI education is also a component of Computer Science for All, the President’s initiative to empower all American students from kindergarten through high school to learn computer science and be equipped with the computational thinking skills they need to be creators, not just consumers, in the digital economy, and to be active citizens in a technology-driven world. The American economy is rapidly shifting, and both educators and business leaders are increasingly recognizing that computer science (CS) is a “new basic” skill necessary for economic opportunity and social mobility. CS for All builds on efforts already being led by parents, teachers, school districts, states, and private sector leaders from across the country and is one way to meet the challenge of preparing a future workforce for the needs of an AI-driven economy.

The Diversity Challenge

All sectors face the challenge of how to diversify the AI workforce. The lack of gender and racial diversity in the AI workforce mirrors the lack of diversity in the technology industry and the field of computer science generally. Unlocking the full potential of the American people, especially in STEM fields, in entrepreneurship, and in the technology industry is a priority of this Administration. The importance of including individuals from diverse backgrounds, experiences, and identities, especially women and members of racial and ethnic groups traditionally underrepresented in STEM, is one of the most critical and high-priority challenges for computer science and AI.

Just 18 percent of computer science graduates today are women, down from a peak of 37 percent in 1984. Though there is a lack of consistently-reported demographic data on the AI workforce, some statistics are available. At the Neural Information Processing Systems (NIPS) Conference in 2015—one of the year’s largest conferences on AI research—just 13.7 percent of conference participants were women.

female. After seeing similarly low representation at a machine intelligence conference, at which she was the only female speaker from industry, the CEO and Co-Founder of Textio, a startup that applies AI to the text of job postings and recruiting emails, decided to further investigate recruitment language in the industry. When the company analyzed 78,768 engineering job listings, they found that job postings for software engineers in the machine intelligence sector had a gender-bias score in favor of men more than twice as high as any other sector.

The diversity challenge is not limited to gender. African Americans, Hispanics, and members of other racial and ethnic minority groups are severely underrepresented, compared to their shares of the U.S. population, in the STEM workforce, in computer science, and in the technology industry workforce, including in the field of AI.

Many of the comments submitted to the OSTP RFI discussed the diversity challenge. Commenters focused on the importance of AI being produced by and for diverse populations. Doing so helps to avoid the negative consequences of narrowly focused AI development, including the risk of biases in developing algorithms, by taking advantage of a broader spectrum of experience, backgrounds, and opinions. These topics were also covered extensively during the public workshops. There is some research on the effects of a lack of diversity in the AI workforce on AI technology design and on the societal impacts of AI. This rich body of research is growing but still lagging behind the literature on broader AI workforce development needs. More research would be beneficial.

**Recommendation 14:** The NSTC Subcommittees on MLAI and NITRD, in conjunction with the NSTC Committee on Science, Technology, Engineering, and Education (CoSTEM), should initiate a study on the AI workforce pipeline in order to develop actions that ensure an appropriate increase in the size, quality, and diversity of the workforce, including AI researchers, specialists, and users.

---


59 The next three high-scoring sectors were back-end engineering, full-stack engineering, and general software engineering. See more: https://textio.ai/gendered-language-in-your-job-post-predicts-the-gender-of-the-person-youll-hire-cd150452407d#.rht0s16ov.
AI, Automation, and the Economy

AI’s central economic effect in the short term will be the automation of tasks that could not be automated before. There is some historical precedent for waves of new automation from which we can learn, and some ways in which AI will be different. Government must understand the potential impacts so it can put in place policies and institutions that will support the benefits of AI, while mitigating the costs.60

Like past waves of innovation, AI will create both benefits and costs. The primary benefit of previous waves of automation has been productivity growth; today’s wave of automation is no different. For example, a 2015 study of robots in 17 countries found that they added an estimated 0.4 percentage point on average to those countries’ annual GDP growth between 1993 and 2007, accounting for just over one-tenth of those countries’ overall GDP growth during that time.61

One important concern arising from prior waves of automation, however, is the potential impact on certain types of jobs and sectors, and the resulting impacts on income inequality. Because AI has the potential to eliminate or drive down wages of some jobs, especially low- and medium-skill jobs, policy interventions will likely be needed to ensure that AI’s economic benefits are broadly shared and that inequality is diminished and not worsened as a consequence.

The economic policy questions raised by AI-driven automation are important but they are best addressed by a separate White House working group. The White House will conduct an additional interagency study on the economic impact of automation on the economy and recommended policy responses, to be published in the coming months.

Recommendation 15: The Executive Office of the President should publish a follow-on report by the end of this year, to further investigate the effects of AI and automation on the U.S. job market, and outline recommended policy responses.

60 Jason Furman, “Is This Time Different? The Opportunities and Challenges of Artificial Intelligence.”

Fairness, Safety, and Governance

As AI technologies gain broader deployment, technical experts and policy analysts have raised concerns about unintended consequences. The use of AI to make consequential decisions about people, often replacing decisions made by human actors and institutions, leads to concerns about how to ensure justice, fairness, and accountability—the same concerns voiced previously in the “Big Data” context.62 The use of AI to control physical-world equipment leads to concerns about safety, especially as systems are exposed to the full complexity of the human environment.

At a technical level, the challenges of fairness and safety are related. In both cases, practitioners strive to prevent intentional discrimination or failure, to avoid unintended consequences, and to generate the evidence needed to give stakeholders justified confidence that unintended failures are unlikely.

Justice, Fairness, and Accountability

A common theme in the Law and Governance, AI for Social Good, and Social and Economic Impacts workshops was the need to ensure that AI promotes justice and fairness, and that AI-based processes are accountable to stakeholders. This issue was highlighted previously in the Administration’s first Big Data report63 published in May 2014, and the follow-up report on Big Data, Algorithmic Systems, Opportunity, and Civil Rights,64 published in May 2016.

In the criminal justice system, some of the biggest concerns with Big Data are the lack of data and the lack of quality data.65 AI needs good data. If the data is incomplete or biased, AI can exacerbate problems of bias. It is important that anyone using AI in the criminal justice context is aware of the limitations of current data.

A commonly cited example at the workshops is the use of apparently biased “risk prediction” tools by some judges in criminal sentencing and bail hearings as well as by some prison officials in assignment and parole decisions, as detailed in an extensively researched ProPublica article.66 The article presented evidence suggesting that a commercial risk scoring tool used by some judges generates racially biased risk scores. A separate report from Upturn questioned the fairness and efficacy of some predictive policing tools.67

---


Similar issues could impact hiring practices. If a machine learning model is used to screen job applicants, and if the data used to train the model reflects past decisions that are biased, the result could be to perpetuate past bias. For example, looking for candidates who resemble past hires may bias a system toward hiring more people like those already on a team, rather than considering the best candidates across the full diversity of potential applicants.

In response to these concerns, several workshop speakers argued for greater transparency when AI tools are used for public purposes. One speaker compared the role of AI to the role of administrative agencies in public decision-making. Authority is delegated to an agency due to the agency’s subject-matter expertise, but the delegation is constrained by due process protections, measures promoting transparency and oversight, and limits on the scope of the delegated authority. Some speakers called for the development of an analogous theory of how to maintain accountability when delegating decision-making power to machines. Transparency concerns focused not only on the data and algorithms used, but also on the potential to have some form of explanation for any AI-based determination.

At the same workshops, AI experts cautioned that there are inherent challenges in trying to understand, predict, and explain the behavior of advanced AI systems, due to the complexity of the systems and the large volume of data they use.

The difficulty of understanding machine learning results is at odds with the common misconception that complex algorithms always do what their designers choose to have them do, and therefore that bias will creep into an algorithm if and only if its developers themselves suffer from conscious or unconscious bias. It is certainly true that a technology developer who wants to produce a biased algorithm can do so, and that unconscious bias may cause practitioners to apply insufficient effort to preventing bias. In practice, however, unbiased developers with the best intentions can inadvertently produce systems with biased results, because even the developers of an AI system may not understand it well enough to prevent unintended outcomes.

Moritz Hardt suggested an illustrative example of how bias might emerge unintentionally from the machine learning process. He postulated a machine learning model trained to distinguish people’s real names from false names. The model might determine that a name is more likely to be false if the first-name part of it is unique in the data set. This rule might have predictive power across the whole population, because false names are more likely to be fanciful and therefore unique. However, if there is an ethnic group that is a small minority of the population and tends to use a different set of first names than the majority population, these distinctive names are more likely to be unique in the sample, and therefore more likely to be incorrectly classified as false names. This effect would arise not because of any special treatment of the minority group’s names, and not because the input data is unrepresentative of the overall population, but simply because the minority group is less numerous.

Andrew Moore, the Dean of Computer Science at Carnegie Mellon University, offered a perspective on the challenge of AI and unforeseen consequences at the workshop on AI Technology, Safety, and Control.

---


69 Some online services require that users sign up for accounts using their real names. Some such services use AI models to detect names suspected of being false, in order to cancel the associated accounts. In such a system, a user whose name is incorrectly classified as false may be unable to sign up for an account, or may have their account canceled unexpectedly.

70 Hardt points to another way that disparate impact may occur. ML models typically become more accurate as the number of examples in the training set increases. In some circumstances, this may cause prediction to be more accurate for a majority group than for a minority. Again, this disparity arises simply because the majority group is more numerous, even if the dataset is representative of the population.
He argued that today, because of the opacity of AI algorithms, the most effective way to minimize the risk of unintended outcomes is through extensive testing—essentially to make a long list of the types of bad outcomes that could occur, and to rule out these outcomes by creating many specialized tests to look for them.

An example of what can go wrong in the absence of extensive testing comes from a trained model for automatically captioning photos, which infamously put the caption “gorilla” on some close-up photos of dark-skinned human faces. This was antithetical to the developers’ values, and it occurred despite testing that showed the model produced accurate results on a high percentage of all photos. These particular errors, although rare, had negative consequences that were beyond the understanding of the model, which had no built-in concept of race, nor any understanding of the relevant historical context. One way to prevent this type of error would have involved extensive testing of the algorithm to scrutinize how human faces, in particular, are labeled, including examination of some results by people who could recognize unacceptable outcomes that the model wouldn’t catch.

Ethical training for AI practitioners and students is a necessary part of the solution. Ideally, every student learning AI, computer science, or data science would be exposed to curriculum and discussion on related ethics and security topics. However, ethics alone is not sufficient. Ethics can help practitioners understand their responsibilities to all stakeholders, but ethical training needs to be augmented with the technical capability to put good intentions into practice by taking technical precautions as a system is built and tested.

As practitioners strive to make AI systems more just, fair and accountable, there are often opportunities to make technology an aid to accountability rather than a barrier to it. Research to improve the interpretability of machine learning results is one example. Having an interpretable model that helps people understand a decision empowers them to interrogate the assumptions and processes behind it.

There are several technical approaches to enhancing the accountability and robustness of complex algorithmic decisions. A system can be tested “in the wild” by presenting it with situations and observing its behavior. A system can be subjected to black-box testing, in which it is presented with synthetic inputs and its behavior is observed, enabling behavior to be tested in scenarios that might not occur naturally. Some or all of the technical details of a system’s design can be published, enabling analysts to replicate it and analyze aspects of its internal behavior that might be difficult to characterize with testing alone. In some cases it is possible to publish information that helps the public evaluate a system’s risk of bias, while withholding other information about the system as proprietary or private.

**Safety and Control**

At the workshops, AI experts said that one of the main factors limiting the deployment of AI in the real world is concern about safety and control. If practitioners cannot achieve justified confidence that a system is safe and controllable, so that deploying the system does not create an unacceptable risk of serious negative consequences, then the system cannot and should not be deployed.

---

71 Some institutions may choose to incorporate ethics into existing courses. Others may choose to introduce separate courses on ethics.

72 Black-box testing allows a system to be presented with fictionalized data, which enables comprehensive experiments that vary individual attributes of an individual as well as larger numbers of experiments than might be possible for in-the-wild testing. See, e.g., Anupam Datta, Shayak Sen, and Yair Zick, “Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems,” *Proceedings of 37th IEEE Symposium on Security and Privacy*, 2016.
A major challenge in safety and control is building systems that can safely transition from the “closed world” of the laboratory into the outside “open world” where unpredictable things can happen. In the open world, a system is likely to encounter objects and situations that were not anticipated when it was designed and built. Adapting gracefully to unforeseen situations is difficult yet necessary for safe operation.

On the topic of safety and predictability in AI, several speakers referenced a recent paper entitled “Concrete Problems in AI Safety,” and the first author of the paper spoke at the workshop on Technology, Safety, and Control. The paper uses a running example of an autonomous robot that does housecleaning. The paper’s overview section gives an extended list of the sorts of practical problems that arise in making such a robot effective and safe, which is quoted here:

**Avoiding Negative Side Effects:** How can we ensure that our cleaning robot will not disturb the environment in negative ways while pursuing its goals, e.g., by knocking over a vase because it can clean faster by doing so? Can we do this without manually specifying everything the robot should not disturb?

**Avoiding Reward Hacking:** How can we ensure that the cleaning robot won’t game its reward function? For example, if we reward the robot for achieving an environment free of messes, it might disable its vision so that it won’t find any messes, or cover over messes with materials it can’t see through, or simply hide when humans are around so they can’t tell it about new types of messes.

**Scalable Oversight:** How can we efficiently ensure that the cleaning robot respects aspects of the objective that are too expensive to be frequently evaluated during training? For instance, it should throw out things that are unlikely to belong to anyone, but put aside things that might belong to someone (it should handle stray candy wrappers differently from stray cellphones). Asking the humans involved whether they lost anything can serve as a check on this, but this check might have to be relatively infrequent—can the robot find a way to do the right thing despite limited information?

**Safe Exploration:** How do we ensure that the cleaning robot doesn’t make exploratory moves with very bad repercussions? For example, the robot should experiment with mopping strategies, but putting a wet mop in an electrical outlet is a very bad idea.

**Robustness to Distributional Shift:** How do we ensure that the cleaning robot recognizes, and behaves robustly, when in an environment different from its training environment? For example, heuristics it learned for cleaning factory work floors may be outright dangerous in an office.

These examples illustrate how the “intelligence” of an AI system can be deep but narrow: the system might have a superhuman ability to detect dirt and optimize its mopping strategy, yet not know to avoid swiping a wet mop over an electrical outlet. One way to describe this overall problem is: how can we give intelligent machines common sense? Researchers are making slow progress on these sorts of problems.

**AI Safety Engineering**

A common theme at the Technology, Safety, and Control workshop was the need to connect open-world AI methods with the broader field of safety engineering. Experience in building other types of safety-critical systems, such as aircraft, power plants, bridges, and vehicles, has much to teach AI practitioners about verification and validation, how to build a safety case for a technology, how to manage risk, and how to communicate with stakeholders about risk.

---

At present, the practice of AI, especially in fast-moving areas of machine learning, can be as much art as science. Certain aspects of practice are not backed by a well-developed theory but instead rely on intuitive judgment and experimentation by practitioners. This is not unusual in newly emerging areas of technology, but it does limit the application of the technology in practice. Some stakeholders have suggested a need to grow AI into a more mature engineering field.

As engineering fields mature, they typically move from an initial “craft” stage characterized by intuition-driven creation by talented amateurs and a do-it-yourself spirit; to a second commercial stage involving skilled practitioners, pragmatic improvement, widely accepted rules-of-thumb, and organized manufacture for sale; to a mature stage that integrates more rigorous methods, educated professionals, well-established theory, and greater specialization of products. Most engineering fields, having a much longer history than modern AI, have reached a mature stage.

In general, mature engineering fields have greater success in creating systems that are predictable, reliable, robust, safe, and secure. Continuing the progress toward AI becoming a mature engineering field will be one of the key enablers of safety and controllability as more complex systems are built.

Recommendation 16: Federal agencies that use AI-based systems to make or provide decision support for consequential decisions about individuals should take extra care to ensure the efficacy and fairness of those systems, based on evidence-based verification and validation.

Recommendation 17: Federal agencies that make grants to state and local governments in support of the use of AI-based systems to make consequential decisions about individuals should review the terms of grants to ensure that AI-based products or services purchased with Federal grant funds produce results in a sufficiently transparent fashion and are supported by evidence of efficacy and fairness.

Recommendation 18: Schools and universities should include ethics, and related topics in security, privacy, and safety, as an integral part of curricula on AI, machine learning, computer science, and data science.

Recommendation 19: AI professionals, safety professionals, and their professional societies should work together to continue progress toward a mature field of AI safety engineering.

74 See, e.g., Mary Shaw, Prospects for an Engineering Discipline of Software, IEEE Software 7(6), November 1990.
Global Considerations and Security

In addition to the long-term challenges of AI and the specific issues relating to fairness and safety, AI poses consequential policy questions in international relations, cybersecurity, and defense.

International Cooperation

AI has been a topic of interest in recent international discussions as countries, multilateral institutions, and other stakeholders have begun to assess the benefits and challenges of AI. Dialogue and cooperation between these entities could help advance AI R&D and harness AI for good, while also addressing pertinent challenges. In particular, several breakthroughs in AI are the direct or indirect result of collaborative research involving people, resources, and institutions in multiple countries. As with other digital policies, countries will need to work together to identify opportunities for cooperation and develop international frameworks that will help promote AI R&D and address any challenges. The United States, a leader in AI R&D, can continue to play a key role in global research coordination through government-to-government dialogues and partnerships.

International engagement is necessary to fully explore the applications of AI in health care, automation in manufacturing, and information and communication technologies (ICTs). AI applications also have the potential to address global issues such as disaster preparedness and response, climate change, wildlife trafficking, the digital divide, jobs, and smart cities. The State Department foresees privacy concerns, safety of autonomous vehicles, and AI’s impact on long-term employment trends as AI-related policy areas to watch in the international context.

In support of U.S. foreign policy priorities in this space—including ensuring U.S. international leadership and economic competitiveness—the U.S. Government has engaged on AI R&D and policy issues in bilateral discussions with other countries, including Japan, the Republic of Korea, Germany, Poland, the United Kingdom, and Italy, as well as in multilateral fora. International AI policy issues and the economic impacts of AI have also been raised in the UN, the G-7, the Organization for Economic Cooperation and Development (OECD), and the Asia-Pacific Economic Cooperation (APEC). The U.S. Government expects AI to be a topic of increasing interest in international engagements.

The United States has been committed to working with industry and relevant standards organizations, in order to facilitate the development of international standards in a manner that is industry-led; voluntary; consensus-driven; and based on principles of transparency, openness, and market needs. The U.S. approach is formalized in law (NTTAA, PL 104-113) and policy (OMB Circular A-119) and reiterated in the United States Standards Strategy.

Recommendation 20: The U.S. Government should develop a government-wide strategy on international engagement related to AI, and develop a list of AI topical areas that need international engagement and monitoring.

Recommendation 21: The U.S. Government should deepen its engagement with key international stakeholders, including foreign governments, international organizations, industry, academia, and others, to exchange information and facilitate collaboration on AI R&D.

AI and Cybersecurity

Today’s Narrow AI has important applications in cybersecurity, and is expected to play an increasing role for both defensive (reactive) measures and offensive (proactive) measures.

Currently, designing and operating secure systems requires a large investment of time and attention from experts. Automating this expert work, partially or entirely, may enable strong security across a much broader range of systems and applications at dramatically lower cost, and may increase the agility of cyber defenses. Using AI may help maintain the rapid response required to detect and react to the landscape of ever evolving cyber threats. There are many opportunities for AI and specifically machine learning systems to help cope with the sheer complexity of cyberspace and support effective human decision making in response to cyberattacks.

Future AI systems could perform predictive analytics to anticipate cyberattacks by generating dynamic threat models from available data sources that are voluminous, ever-changing, and often incomplete. These data include the topology and state of network nodes, links, equipment, architecture, protocols, and networks. AI may be the most effective approach to interpreting these data, proactively identifying vulnerabilities, and taking action to prevent or mitigate future attacks.

Results to-date in DARPA’s Cyber Grand Challenge (CGC) competition demonstrate the potential of this approach. The CGC was designed to accelerate the development of advanced, autonomous systems that can detect, evaluate, and patch software vulnerabilities before adversaries have a chance to exploit them. The CGC Final Event was held on August 4, 2016. To fuel follow-on research and parallel competition, all of the code produced by the automated systems during the CGC Final Event has been released as open source to allow others to reverse engineer it and learn from it.

AI systems also have their own cybersecurity needs. AI-driven applications should implement sound cybersecurity controls to ensure integrity of data and functionality, protect privacy and confidentiality, and maintain availability. The recent Federal Cybersecurity R&D Strategic Plan highlighted the need for “sustainably secure systems development and operation.” Advances in cybersecurity will be critical in making AI solutions secure and resilient against malicious cyber activities, particularly as the volume and type of tasks conducted by governments and private sector businesses using Narrow AI increases.

Finally, AI could support planning, coordinating, integrating, synchronizing, and directing activities to operate and defend U.S. government networks and systems effectively, provide assistance in support of secure operation of private-sector networks and systems, and enable action in accordance with all applicable laws, regulations and treaties.

76 https://www.cybergrandchallenge.com

AI in Weapon Systems

The United States has incorporated autonomy into certain weapon systems for decades. These technological improvements may allow for greater precision in the use of these weapon systems and safer, more humane military operations. Precision-guided munitions allow an operation to be completed with fewer weapons expended and with less collateral damage, and remotely-piloted vehicles can lessen the risk to military personnel by placing greater distance between them and danger. Nonetheless, moving away from direct human control of weapon systems involves some risks and can raise legal and ethical questions. The key to further incorporating autonomous and semi-autonomous weapon systems into U.S. defense planning and force structure is to continue ensuring that all our weapon systems, including autonomous weapon systems, are being used in a manner consistent with international humanitarian law. In addition, the U.S. Government should continue taking appropriate steps to control proliferation, and working with partners and Allies to develop standards related to the development and use of such weapon systems.

Over the past several years, in particular, issues concerning the development of so-called “Lethal Autonomous Weapon Systems” (LAWS) have been raised by technical experts, ethicists, and others in the international community. The United States has actively participated in the ongoing international discussion on LAWS in the context of the Convention on Certain Conventional Weapons (CCW), and anticipates continued robust international discussion of these potential weapon systems going forward.

State Parties to the CCW are discussing technical, legal, military, ethical, and other issues involved with emerging technologies, although it is clear that there is no common understanding of LAWS. Some States have conflated LAWS with remotely piloted aircraft (military “drones”), a position which the United States opposes, as remotely-piloted craft are, by definition, directly controlled by humans just as manned aircraft are. Other States have focused on artificial intelligence, robot armies, or whether “meaningful human control” – an undefined term – is exercised over life-and-death decisions. The U.S. priority has been to reiterate that all weapon systems, autonomous or otherwise, must adhere to international

---


79 See, e.g., DeepMind comment, Human Rights Watch comment to the OSTP Request for Information on Artificial Intelligence.

humanitarian law, including the principles of distinction\(^\text{81}\) and proportionality.\(^\text{82}\) For this reason, the United States has consistently noted the importance of the weapons review process in the development and adoption of new weapon systems. The CCW will decide on whether and how to conduct future meetings on LAWS and associated issues during its Review Conference in December 2016.

The U.S. government is also conducting a comprehensive review of the implications of autonomy in defense systems. In November 2012, the Department of Defense (DoD) issued DoD Directive 3000.09, “Autonomy in Weapon Systems,” which outlines requirements for the development and fielding of autonomous and semi-autonomous weapons. Weapon systems capable of autonomously selecting and engaging targets with lethal force require senior-level DoD reviews and approval before those weapon systems enter formal development and again before fielding. The DoD Directive neither prohibits nor encourages such development, but requires it to proceed carefully and only after review and approval by senior defense officials. Among other things, the DoD Directive requires that autonomous and semi-autonomous weapon systems are rigorously tested and that personnel are trained appropriately in their use to advance international norms pertaining to armed conflict.

AI has the potential to provide significant benefits across a range of defense-related activities. Non-lethal activities such as logistics, maintenance, base operations, veterans’ healthcare, lifesaving battlefield medical assistance and casualty evacuation, personnel management, navigation, communication, cyber-defense, and intelligence analysis can benefit from AI, making American forces safer and more effective. AI may also play an important role in new systems for protecting people and high-value fixed assets and deterring attacks through non-lethal means. Ultimately, these applications may turn out to be the most important for DoD.

Given advances in military technology and artificial intelligence more broadly, scientists, strategists, and military experts all agree that the future of LAWS is difficult to predict and the pace of change is rapid. Many new capabilities may soon be possible, and quickly able to be developed and operationalized. The Administration is engaged in active, ongoing interagency discussions to work toward a government-wide policy on autonomous weapons consistent with shared human values, national security interests, and international and domestic obligations.

\textit{Recommendation 23: The U.S. Government should complete the development of a single, government-wide policy, consistent with international humanitarian law, on autonomous and semi-autonomous weapons.}\n
\(^{81}\) The principle of distinction requires parties to a conflict to distinguish between the civilian population and combatants and between civilian objects and military objectives, and to direct their operations only against military objectives.

\(^{82}\) The principle of proportionality prohibits attacks that may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.
Conclusion

AI can be a major driver of economic growth and social progress, if industry, civil society, government, and the public work together to support development of the technology, with thoughtful attention to its potential and to managing its risks.

Government has several roles to play. It should convene conversations about important issues and help to set the agenda for public debate. It should monitor the safety and fairness of applications as they develop, and adapt regulatory frameworks to encourage innovation while protecting the public. It should support basic research and the application of AI to public goods, as well as the development of a skilled, diverse workforce. And government should use AI itself, to serve the public faster, more effectively, and at lower cost.

Many areas of public policy, from education and the economic safety net, to defense, environmental preservation, and criminal justice, will see new opportunities and new challenges driven by the continued progress of AI. Government must continue to build its capacity to understand and adapt to these changes.

As the technology of AI continues to develop, practitioners must ensure that AI-enabled systems are governable; that they are open, transparent, and understandable; that they can work effectively with people; and that their operation will remain consistent with human values and aspirations. Researchers and practitioners have increased their attention to these challenges, and should continue to focus on them.

Developing and studying machine intelligence can help us better understand and appreciate our human intelligence. Used thoughtfully, AI can augment our intelligence, helping us chart a better and wiser path forward.
Recommendations in this Report

This section collects all of the recommendations in this report, for ease of reference.

**Recommendation 1: Private and public institutions are encouraged to examine whether and how they can responsibly leverage AI and machine learning in ways that will benefit society.** Social justice and public policy institutions that do not typically engage with advanced technologies and data science in their work should consider partnerships with AI researchers and practitioners that can help apply AI tactics to the broad social problems these institutions already address in other ways.

**Recommendation 2: Federal agencies should prioritize open training data and open data standards in AI.** The government should emphasize the release of datasets that enable the use of AI to address social challenges. Potential steps may include developing an “Open Data for AI” initiative with the objective of releasing a significant number of government data sets to accelerate AI research and galvanize the use of open data standards and best practices across government, academia, and the private sector.

**Recommendation 3: The Federal Government should explore ways to improve the capacity of key agencies to apply AI to their missions.** For example, Federal agencies should explore the potential to create DARPA-like organizations to support high-risk, high-reward AI research and its application, much as the Department of Education has done through its proposal to create an “ARPA-ED,” to support R&D to determine whether AI and other technologies could significantly improve student learning outcomes.

**Recommendation 4: The NSTC MLAI subcommittee should develop a community of practice for AI practitioners across government.** Agencies should work together to develop and share standards and best practices around the use of AI in government operations. Agencies should ensure that Federal employee training programs include relevant AI opportunities.

**Recommendation 5: Agencies should draw on appropriate technical expertise at the senior level when setting regulatory policy for AI-enabled products.** Effective regulation of AI-enabled products requires collaboration between agency leadership, staff knowledgeable about the existing regulatory framework and regulatory practices generally, and technical experts with knowledge of AI. Agency leadership should take steps to recruit the necessary technical talent, or identify it in existing agency staff, and should ensure that there are sufficient technical “seats at the table” in regulatory policy discussions.

**Recommendation 6: Agencies should use the full range of personnel assignment and exchange models (e.g. hiring authorities) to foster a Federal workforce with more diverse perspectives on the current state of technology.**

**Recommendation 7: The Department of Transportation should work with industry and researchers on ways to increase sharing of data for safety, research, and other purposes.** The future roles of AI in surface and air transportation are undeniable. Accordingly, Federal actors should focus in the near-term on developing increasingly rich sets of data, consistent with consumer privacy, that can better inform policy-making as these technologies mature.
Recommendation 8: The U.S. Government should invest in developing and implementing an advanced and automated air traffic management system that is highly scalable, and can fully accommodate autonomous and piloted aircraft alike.

Recommendation 9: The Department of Transportation should continue to develop an evolving framework for regulation to enable the safe integration of fully automated vehicles and UAS, including novel vehicle designs, into the transportation system.

Recommendation 10: The NSTC Subcommittee on Machine Learning and Artificial Intelligence should monitor developments in AI, and report regularly to senior Administration leadership about the status of AI, especially with regard to milestones. The Subcommittee should update the list of milestones as knowledge advances and the consensus of experts changes over time. The Subcommittee should consider reporting to the public on AI developments, when appropriate.

Recommendation 11: The Government should monitor the state of AI in other countries, especially with respect to milestones.

Recommendation 12: Industry should work with government to keep government updated on the general progress of AI in industry, including the likelihood of milestones being reached soon.

Recommendation 13: The Federal government should prioritize basic and long-term AI research. The Nation as a whole would benefit from a steady increase in Federal and private-sector AI R&D, with a particular emphasis on basic research and long-term, high-risk research initiatives. Because basic and long-term research especially are areas where the private sector is not likely to invest, Federal investments will be important for R&D in these areas.

Recommendation 14: The NSTC Subcommittees on MLAI and NITRD, in conjunction with the NSTC Committee on Science, Technology, Engineering, and Education (CoSTEM), should initiate a study on the AI workforce pipeline in order to develop actions that ensure an appropriate increase in the size, quality, and diversity of the workforce, including AI researchers, specialists, and users.

Recommendation 15: The Executive Office of the President should publish a follow-on report by the end of this year, to further investigate the effects of AI and automation on the U.S. job market, and outline recommended policy responses.

Recommendation 16: Federal agencies that use AI-based systems to make or provide decision support for consequential decisions about individuals should take extra care to ensure the efficacy and fairness of those systems, based on evidence-based verification and validation.

Recommendation 17: Federal agencies that make grants to state and local governments in support of the use of AI-based systems to make consequential decisions about individuals should review the terms of grants to ensure that AI-based products or services purchased with Federal grant funds produce results in a sufficiently transparent fashion and are supported by evidence of efficacy and fairness.
Recommendation 18: Schools and universities should include ethics, and related topics in security, privacy, and safety, as an integral part of curricula on AI, machine learning, computer science, and data science.

Recommendation 19: AI professionals, safety professionals, and their professional societies should work together to continue progress toward a mature field of AI safety engineering.

Recommendation 20: The U.S. Government should develop a government-wide strategy on international engagement related to AI, and develop a list of AI topical areas that need international engagement and monitoring.

Recommendation 21: The U.S. Government should deepen its engagement with key international stakeholders, including foreign governments, international organizations, industry, academia, and others, to exchange information and facilitate collaboration on AI R&D.

Recommendation 22: Agencies’ plans and strategies should account for the influence of AI on cybersecurity, and of cybersecurity on AI. Agencies involved in AI issues should engage their U.S. Government and private-sector cybersecurity colleagues for input on how to ensure that AI systems and ecosystems are secure and resilient to intelligent adversaries. Agencies involved in cybersecurity issues should engage their U.S. Government and private sector AI colleagues for innovative ways to apply AI for effective and efficient cybersecurity.

Recommendation 23: The U.S. Government should complete the development of a single, government-wide policy, consistent with international humanitarian law, on autonomous and semi-autonomous weapons.
Acronyms

AAAI  Association for the Advancement of Artificial Intelligence
AGI  Artificial General Intelligence
AI  Artificial Intelligence
APEC  Asia-Pacific Economic Cooperation
BRAIN  Brain Research through Advancing Innovative Neurotechnologies
CALO  Cognitive Agent that Learns and Organizes
CCC  Computing Community Consortium
CCW  Convention on Certain Conventional Weapons
CEA  Council of Economic Advisers
CEO  Chief Executive Officer
CGC  Cyber Grand Challenge (run by DARPA)
CoSTEM  Committee on Science Technology, Engineering, and Education (component of NSTC)
CS  Computer Science
DARPA  Defense Advanced Research Projects Agency
DoD  Department of Defense
DOT  Department of Transportation
FAA  Federal Aviation Administration
FMVSS  Federal Motor Vehicle Safety Standards
IARPA  Intelligence Advanced Research Projects Activity
ICTs  Information and Communication Technologies
IPA  Intergovernmental Personnel Act
LAWS  Lethal Autonomous Weapon Systems
MLAI  Machine Learning and Artificial Intelligence (subcommittee of NSTC)
NAS  National Airspace System
NEC  National Economic Council
NHTSA  National Highway Traffic Safety Administration
NIH  National Institutes of Health
NIPS  Neural Information Processing Systems conference
NITRD  Networking and Information Technology Research and Development (subcommittee of NSTC)
NSF  National Science Foundation
NSTC  National Science and Technology Council
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD</td>
<td>Organization for Economic Cooperation and Development</td>
</tr>
<tr>
<td>OMB</td>
<td>Office of Management and Budget</td>
</tr>
<tr>
<td>ONR</td>
<td>Office of Naval Research</td>
</tr>
<tr>
<td>OSTP</td>
<td>Office of Science and Technology Policy</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Research and Development</td>
</tr>
<tr>
<td>RFI</td>
<td>Request For Information</td>
</tr>
<tr>
<td>STEM</td>
<td>Science, Technology, Engineering, and Mathematics</td>
</tr>
<tr>
<td>UAS</td>
<td>Unmanned Aerial System</td>
</tr>
<tr>
<td>UTM</td>
<td>UAS Traffic Management</td>
</tr>
</tbody>
</table>
References


PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE


Mary Shaw, Prospects for an Engineering Discipline of Software, IEEE Software 7(6), November 1990.


Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian,


As companies building the technologies in Artificial Intelligence and Robotics that may be repurposed to develop autonomous weapons, we feel especially responsible in raising this alarm. We warmly welcome the decision of the UN's Conference of the Convention on Certain Conventional Weapons (CCW) to establish a Group of Governmental Experts (GGE) on Lethal Autonomous Weapon Systems. Many of our researchers and engineers are eager to offer technical advice to your deliberations.

We commend the appointment of Ambassador Amandeep Singh Gill of India as chair of the GGE. We entreat the High Contracting Parties participating in the GGE to work hard at finding means to prevent an arms race in these weapons, to protect civilians from their misuse, and to avoid the destabilizing effects of these technologies. We regret that the GGE's first meeting, which was due to start today (August 21, 2017), has been cancelled due to a small number of states failing to pay their financial contributions to the UN. We urge the High Contracting Parties therefore to double their efforts at the first meeting of the GGE now planned for November.

Lethal autonomous weapons threaten to become the third revolution in warfare. Once developed, they will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend. These can be weapons of terror, weapons that despot and terrorists use against innocent populations, and weapons hacked to behave in undesirable ways. We do not have long to act. Once this Pandora's box is opened, it will be hard to close. We therefore implore the High Contracting Parties to find a way to protect us all from these dangers.

Translations: [Chinese](http://futureoflife.org/open-letter-united-nations-convention-certain-conventional-weapons-chinese)
FULL LIST OF SIGNATORIES TO THE OPEN LETTER

To add your company, please contact Toby Walsh at tw@cse.unsw.edu.au.

Tiberio Caetano, founder & Chief Scientist at Ambiata, Australia.
Mark Chatterton and Leo Gui, founders, MD & of Ingenious AI, Australia.
Charles Gretton, founder of Hivery, Australia.
Brad Lorge, founder & CEO of Premonition.io, Australia.
Brenton O’Brien, founder & CEO of Microbric, Australia.
Samir Sinha, founder & CEO of Robonomics AI, Australia.
Ivan Storr, founder & CEO, Blue Ocean Robotics, Australia.
Peter Turner, founder & MD of Tribotix, Australia.
Yoshua Bengio, founder of Element AI & Montreal Institute for Learning Algorithms, Canada.
Ryan Gariepy, founder & CTO, Clearpath Robotics, found & CTO of OTTO Motors, Canada.
Geoffrey Hinton, founder of DNNResearch Inc, Canada.
James Chow, founder & CEO of UBTECH Robotics, China.
Robert Li, founder & CEO of Sankobot, China.
Marek Rosa, founder & CEO of GoodAI, Czech Republic.
Soren Tranberg Hansen, founder & CEO of Brainbots, Denmark.
Markus Järve, founder & CEO of Krakul, Estonia.
Harri Valpola, founder & CTO of ZenRobotics, founder & CEO of Curious AI Company, Finland.
Esben Østergaard, founder & CTO of Universal Robotics, Denmark.
Raul Bravo, founder & CEO of DIBOTICS, France.
Ivan Burdun, founder & President of AIXTREE, France.
Raphael Cherrier, founder & CEO of Qucit, France.
Alain Garnier, founder & CEO of ARISEM (acquired by Thales), founder & CEO of Jamespot, France.
Charles Ollion, founder & Head of Research at Heuritech, France.
Anis Sahbani, founder & CEO of Enova Robotics, France.
Alexandre Vallette, founder of SNIPS & Ants Open Innovation Labs, France.
Marcus Frei, founder & CEO of NEXT.robotics, Germany.
Kirstinn Thorisson, founder & Director of Icelandic Institute for Intelligence Machines, Iceland.
Fahad Azad, founder of Robosoft Systems, India.
Debashis Das, Ashish Tupate & Jerwin Prabhu, founders (incl. CEO) of Bharati Robotics, India.
Pulkit Gaur, founder & CTO of Gridbots Technologies, India.
Pranay Kishore, founder & CEO of Phi Robotics Research, India.
Shahid Memom, founder & CTO of Vanora Robots, India.
Krishnan Nambari & Shahid Memom, founders, CEO & CTO of Vanora Robotics, India.
Achu Wilson, founder & CTO of Sastra Robotics, India.
Neill Gernon, founder & MD of Atrovate, founder of Dublin.AI, Ireland.
Parsa Ghaffari, founder & CEO of Aylien, Ireland.
Alan Holland, founder & CEO of Keelvar Systems, Ireland.
Alessandro Prest, founder & CTO of LogoGrab, Ireland.
Frank Reeves, founder & CEO of Avvio, Ireland.
Alessio Bonietti, founder & CEO of MindIT, Italy.
Angelo Sudano, founder & CTO of ICan Robotics, Italy.
Domenico Talia, founder and R&D Director of DtoK Labs, Italy.
Shigeo Hirose, Michele Guarneri, Paulo Debenest, & Nah Kitano, founders, CEO & Directors of HiBot Corporation, Japan.
Andrejs Vasiljevs, founder and CEO of Tilde, Latvia.
Luis Samahí García González, founder & CEO of QOLbots, Mexico.
Koen Hindriks & Joachim de Greeff, founders, CEO & COO at Interactive Robotics, the Netherlands.
Maja Rudinac, founder and CEO of Robot Care Systems, the Netherlands.
Jaap van Leeuwen, founder and CEO Blue Ocean Robotics Benelux, the Netherlands.
Rob Brouwer, founder and Director of Operatins, Aeronavics, New Zealand.
Philip Solaris, founder and CEO of X-Craf Enterprises, New Zealand.
Dyrkoren Erik, Martin Ludvigsen & Christine Spiten, founders, CEO, CTO & Head of Marketing at BlueEye Robotics, Norway.
Sergii Korneiev, founder & CEO of BaltRobotics, Poland.
Igor Kuznetsov, founder & CEO of NaviRobot, Russian Federation.
Aleksey Yuzhakov & Oleg Kivokurtsev, founders, CEO & COO of Promobot, Russian Federation.
Junyang Woon, founder & CEO, Infinium Robotics, former Branch Head & Naval Warfare Operations Officer, Singapore.
Jasper Horrell, founder of DeepData, South Africa.
Onno Huyser and Mark van Wyk, founders of FlyH2 Aerospace, South Africa.
Toni Ferrate, founder & CEO of RO-BOTICS, Spain.
José Manuel del Río, founder & CEO of Aisoy Robotics, Spain.
Victor Martin, founder & CEO of Macco Robotics, Spain.
Angel Lis Montesinos, founder & CTO of Neuronalbite, Spain.
Timothy Llewellynn, founder & CEO of nViso, Switzerland.
Francesco Mondada, founder of K-Team, Switzerland.
Jurgen Schmidhuber, Faustino Gomez, Jan Koutník, Jonathan Masci & Bas Steunebrink, founders, President & CEO of Nnaisense, Switzerland.
Satish Ramachandran, founder of AROBOT, United Arab Emirates.
Silas Adekunle, founder & CEO of Reach Robotics, UK.
Steve Allpress, founder & CTO of FiveAI, UK.
John Bishop, founder and Director of Tungsten Centre for Intelligent Data Analytis, UK.
Joel Gibbard and Samantha Payne, founders, CEO & COO of Open Bionics, UK.
Nic Greenway, founder of React AI Ltd (Aiseedo), UK.
Daniel Hulme, founder & CEO of Satalia, UK.
Bradley Kieser, founder & Director of SMS Speedway, UK.
Charlie Muirhead & Tabitha Goldstaub, founders & CEO of CognitionX, UK.
Geoff Pegman, founder & MD of R U Robots, UK.
Demis Hassabis & Mustafa Suleyman, founders, CEO & Head of Applied AI, DeepMind, UK.
Donald Szeto, Thomas Stone & Kenneth Chan, founders, CTO, COO & Head of Engineering of PredictionIO, UK.
Antoine Biondeau, founder & CEO of Sentient Technologies, USA.
Steve Cousins, founder & CEO of Savioke, USA.
Brian Gerkey, founder & CEO of Open Source Robotics, USA.
Ryan Hickman & Soohyun Bae, founders, CEO & CTO of TickTock.AI, USA.
John Hobart, founder & CEO of Coria, USA.
Henry Hu, founder & CEO of Cafe X Technologies, USA.
Zaib Husain, founder & CEO of SparkCognitin, USA.
Alfonso Íñiguez, founder & CEO of Swarm Technology, USA.
Kris Kitchen, founder & Chief Data Scientist at Qieon Research, USA.
Justin Lane, founder of Prospecture Simulation, USA.
Gary Marcus, founder & CEO of Geometric Intelligence (acquired by Uber), USA.
Brian Mingus, founder & CTO of Latently, USA.
Mohammad Musa, founder & CEO at Deepen AI, USA.
Elon Musk, founder, CEO & CTO of SpaceX, co-founder & CEO of Tesla Motor, USA.
Rosanna Myers & Dan Corkum, founders, CEO & CTO of Carbon Robotics, USA.
Erik Nieves, founder & CEO of PlusOne Robotics, USA.
Steve Omohundro, founder & President of Possibility Research, USA.
Jeff Orkin, founder & CEO, Giant Otter Technologies, USA.
Greg Phillips, founder & CEO, ThinkIt Data Solutions, USA.
Dan Reuter, founder & CEO of Electric Movement, USA.
Alberto Rizzoli & Simon Edwardsson, founders & CEO of AI Poly, USA.
Dan Rubins, founder & CEO of Legal Robot, USA.
Stuart Russell, founder & VP of Bayesian Logic Inc., USA.
Andrew Schroeder, founder of WeRobotics, USA.
Stanislav Shalunov, founder & CEO of Clostra, USA.
Gabe Sibley & Alex Flint, founders, CEO & CPO of Zippy.ai, USA.
Martin Spencer, founder & CEO of GeckoSystems, USA.
Peter Stone, Mark Ring & Satinder Singh, founders, President/COO, CEO & CTO of Cogitai, USA.
Michael Stuart, founder & CEO of Lucid Holdings, USA.
Madhuri Trivedi, founder & CEO of OrangeHC, USA.
Massimiliano Versace, founder, CEO & President, Neurala Inc, USA.
Reza Zadeh, founder & CEO of Matroid, USA.
ACCOUNTABILITY AND AUTONOMOUS WEAPONS: MUCH ADO ABOUT NOTHING?

Charles J. Dunlap, Jr.*

INTRODUCTION

One of the hottest topics of emerging technologies of war is autonomous weapons. Within the field, legal and ethical questions are as prominent as the technical ones, and two camps seem to have emerged. The first seeks to ban autonomous weapons entirely, and the second recognizes their inevitability and seeks to regulate them.

A leader of the first group (demanding a total ban) has been Human Rights Watch (HRW). In a major paper, entitled Losing Humanity: The Case Against Killer Robots, HRW - aided by Harvard Law School’s International Human Rights Clinic (IHRC) – attempted to establish that fully autonomous weapons “would not be consistent with international humanitarian law and would increase the risk of death or injury to civilians during armed conflict.” Accordingly, the report concluded, a “preemptive prohibition on their development and use is needed.”

However, that effort was rather thoroughly deconstructed by Professor Michael Schmitt in a rebuttal entitled, “Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics.” Schmitt found that a “principal flaw


1 See e.g., Campaign to Stop Killer Robots, http://www.stopkillerrobots.org/ (last visited Mar. 4, 2016).


3 HUMAN RIGHTS WATCH, LOSING HUMANITY: THE CASE AGAINST KILLER ROBOTS, Nov. 2012, at 1, http://www.hrw.org/sites/default/files/reports/arms1112ForUpload_0_0.pdf [hereinafter LOSING HUMANITY]. The paper defines “fully autonomous” weapons as referring to “to both out-of-the-loop weapons and those that allow a human on the loop, but that are effectively out-of-the-loop weapons because the supervision is so limited.” Additionally, the report defines “out-of-the-loop weapons” as being “[r]obots that are capable of selecting targets and delivering force without any human input or interaction.”

4 Id.

5 Michael N. Schmitt, Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics, HARVARD NATIONAL SECURITY JOURNAL FEATURES (Feb. 5, 2013, 2:07 pm),
in the analysis is a blurring of the distinction between international humanitarian law’s prohibitions on weapons *per se* and those on the unlawful use of otherwise lawful weapons.” He went on to convincingly conclude that “that autonomous weapon systems are not unlawful *per se,*” adding that:

> Their autonomy has no direct bearing on the probability they would cause unnecessary suffering or superfluous injury, does not preclude them from being directed at combatants and military objectives, and need not result in their having effects that an attacker cannot control. Individual systems could be developed that would violate these norms, but autonomous weapon systems are not prohibited on this basis *as a category.*

While it may be that the HRW/IHRC’s report was (and is) popular among anti-autonomous non-governmental organization (NGO) weapons groups, it evidently did not make much headway with nation-states because as of 2015 as “many as forty nations are currently developing military robotics.” Still, the Future of Life Institute gained considerable publicity recently when it distributed an “open letter” signed by many artificial intelligence developers (and others) which called for a “ban on offensive autonomous weapons beyond meaningful human control.”

Apparently undeterred, HRW/IHRC renewed its effort with a new paper released in April of 2015 entitled, *Mind the Gap: The Lack of Accountability for Killer Robots.* This report replays many themes from the previous effort – and


6 Id., at

7 Id., at 35.

8 Id.


12 *Autonomous Weapons: an Open Letter from AI & Robotics Researchers,* Future of Life Institute, Jul. 28, 2015, http://futureoflife.org/AI/open_letter_autonomous_weapons (last visited Mar. 4, 2016). Besides the issue as to what phrase “offensive” autonomous weapons would encompass, it is also not explained as to why non-offensive autonomous weapons “beyond meaningful human control” would not likewise be objectionable.

comes up with the same outright ban recommendation - but alters its tactic towards that end somewhat. This time it seems that autonomous weapons should be banned because, they claim, “neither criminal law nor civil law guarantees adequate accountability for individuals directly or indirectly involved in the use of fully autonomous systems.”

This purpose of this essay is to briefly examine Mind the Gap to see how it relates to legal actualities as to accountability. It will conclude that it deviates from it in material ways, and finds that this new tactic is even more egregiously flawed than HRW/IHRC’s original approach. In point of fact, although no one can “guarantee” accountability, there are sufficient legal tools to do so when appropriate; autonomous weapons are not somehow exempted from legal regimes applicable to other weapons or the law of war more generally. This essay will contend – as others have – that it is better to develop norms to control these systems than to attempt to ban them outright.

THE THRESHOLD QUESTION: IS PERSONAL ACCOUNTABILITY AN ESSENTIAL ELEMENT OF THE LEGALITY OF A WEAPON?

In presenting its contentions HRW/IHRC seems to confuse the issue of personal accountability with the legality of a weapons system itself. In fact, international law has no such requirement, and Mind the Gap identifies none. Article 36 of Protocol 1 of the Geneva Conventions does call upon parties to “determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law.” The Protocol goes to focus weapons’ legality on whether it can be “directed at a specific military objective,” and prohibits those which are “of a nature to strike military objectives and civilians or civilian objects without distinction.”

For its part, the U.S. (which is not a party to Protocol 1) does not recognize Article 36 as part of customary international law, but conducts such reviews as matter of long standing policy. These reviews follow standard international law

---

14 Id. at 1.
16 Protocol 1, art. 51. The International Committee of the Red Cross considers these prohibitions to be part of customary international law,
in that they ask the following questions:

- whether the weapon’s intended use is calculated to cause superfluous injury;
- whether the weapon is inherently indiscriminate; and
- whether the weapon falls within a class of weapons that has been specifically prohibited.\(^\text{18}\)

There are other possible prohibitions on weapon – such as a specific treaty – but *none* conditions legality on the ability to assign blame to a specific individual. One of the world’s foremost weapons’ law authorities, Professor William Boothby unequivocally dismisses the suggestion that *individual accountability* an essential element of the legality of a weapon. He says:

> The lawfulness of an autonomous weapon system under current international law does not, in my view, turn on the ability or otherwise to fix any identifiable individual with liability in the event of an unsatisfactory attack. Sometimes it will be possible to assign responsibility to an identifiable individual, sometimes it will not.\(^\text{19}\)

This plainly reflects not only *lex lata*, but the practical realities of war. Yet *Mind the Gap* wrongly conflates the imperative under international law to investigate and prosecute “grave breaches” with the separate issue of the legality of a particular weapon.\(^\text{20}\) Many if not most inarguably lawful weapons might still be used in an unlawful manner, but that does not lead to calls for bans. The focus should be on the way the system is used. To be clear, the International Committee of the Red Cross (ICRC) makes the point plainly that “a weapon that *can be used with precision* can also be abusively used against the civilian population. In this case, it is *not the weapon which is prohibited*, but the method or the way in which it is used.”\(^\text{21}\)

**DOES PERSONAL ACCOUNTABILITY REALLY DETER?**

---

\(^{18}\) *Id.*

\(^{19}\) Email from Air Commodore William Boothby, RAF (Ret.), Associate Fellow at the Geneva Centre for Security Policy to the author (Oct. 11, 2015 5:26 PM) (on file with the author). Air Cdre Boothby is the author of *WEAPONS AND THE LAW OF ARMED CONFLICT* (2009).

\(^{20}\) *MIND THE GAP*, at 15-17.

Mind the Gap seeks to support its thesis that individual liability is essential as a matter of international law to a weapon’s legality by lecturing the reader with its interpretation of the “Purposes of Criminal Responsibility.” That rendition is mainly a rather basic recitation of the standard criminal justice themes of deterrence and retribution, as well as what HRW/IHRC calls “compensatory justice.”

However, in the context of international criminal tribunals (ICT) convened to judge atrocities and other grievous offenses against human rights, the utility of personal accountability for the purpose of deterrence is debatable. Terminating conflicts and rebuilding societies after them is a complex task, and efforts to impose individual liability in the name of deterrence against future acts may actually prove to be counterproductive. As two scholars put it, there are “reasons to be wary of the deterrence promise of ICTs”, adding that “it is dangerously naïve to ignore the possibility that ICTs might not only lack any significant deterrence benefits, but might actually exacerbate conflicts in weak states.”

Furthermore, some experts question whether the psychology of war criminals—and particularly that of the most egregious among them—is such that they are even amenable to deterrence under any circumstances. Consider:

Many argue that war crimes tribunals offer no deterrent to potential criminals whatsoever. People with strong convictions against a certain religious or ethnic group will likely not feel any less hatred for that group just because a possible tribunal looms in the future. Both Hitler and Pol Pot believed they would be revered by future generations for the extreme measures they took to change the makeup of their societies. These leaders were inspired by their visions of the future and it is unlikely the prospect of a war crimes tribunal would have swayed either dictator.

In their report, HRW/IHRC never really accounts for the markedly diminished status of deterrence not only in the U.S., but in the international community generally (irrespective of any connection with autonomous weapons). In fact, that community has largely rejected the most coercive of all deterents—the death
penalty – in international tribunals.  

More broadly, even the long-accepted principle of belligerent reprisal – something explicitly aimed at deterring a belligerent from continuing to violate international law – has been eviscerated by Protocol 1’s restrictions on reprisals, even against objects. This is despite the fact that, as Professor Michael Newton argues, “[r]easonable reprisals grounded on an empirical assessment of their deterrent value or framed as appropriate punishment for prior acts of terror may be the most morally acceptable and humane strategy for serving a strategic imperative of civilized society.”

To the extent that HRW/IHRC is underpinning their call for a ban on autonomous weapons based on the supposed indispensability of personal deterrence, that case has yet to be made. Given the political reality that individual criminal accountability does not always serve the strategic need for societal reconciliation, as well as the fact the international community has progressively deconstructed the tools of traditional deterrence, it is unclear that there is any significant contemporary norm to support HRW/IHRC suppositions as to deterrence, qua deterrence, and the propriety of their proposed ban on autonomous weapons.

**WHY DOES HRW/IHRC THINK ACCOUNTABILITY CANNOT BE ACHIEVED?**

Perhaps the most puzzling part of *Mind the Gap* is their central thesis: that there is not – and could not be – accountability for illicit use of autonomous weapons. Inexplicably, they believe it is necessary to make the rather obvious observation that “robots are not men” and go on to conclude in what we must assume is a serious syllogism, that “fully autonomous weapons could not have the mental state to make these wrongful actions crimes.” This, of course, is irrelevant as international law has never sought to impose criminal liability on weapons themselves.

---

30 *MIND THE GAP*, at 15-17.
31 Interestingly, in her book *LIABILITY FOR CRIMES INVOLVING ARTIFICIAL INTELLIGENCE SYSTEMS* (2015) Gabriel Hallevy argues that current criminal law – much developed from criminal liability of corporations – could cause “odd consequences, such as the imposition of criminal liability on machines.” Whatever resonance this may have in domestic law, or even international law outside of the law of war, it does not apply to the lawfulness of weapons.
As Professor Michael Schmitt points out, contentions about accountability have “muddled” the debate about autonomous weapons. As he observes, it is not difficult to map out how it would be allocated:

Clearly, any commander who decides to launch AWS [autonomous weapons system] into a particular environment is, as with any other weapon systems, accountable under international criminal law for that decision. Nor will developers escape accountability if they design systems, autonomous or not, meant to conduct operations that are not IHL [international humanitarian law] compliant. And States can be held accountable under the laws of State responsibility should their armed forces use AWS in an unlawful manner.32

Likewise, Armin Krishnan, concludes in his book, KILLER ROBOTS: LEGALITY AND ETHICALITY OF AUTONOMOUS WEAPONS, that the “legal problems with regard to accountability might be far smaller than some critics of military robots believe.”33 He sensibly points out that if “the robot does not operate within the boundaries of its specified parameters it is the manufacture’s fault.” Similarly, he says that if the robot is “used in circumstances that make its use illegal, then it is the commander’s fault.”34

But Mind the Gap assumes that a “gap could arise because fully autonomous weapons by definition would have capacity to act autonomously and therefore could launch independently and unforeseeably and indiscriminate attack against civilians…”35 Here’s the question: what system of justice in the civilized world attempts to impose criminal liability on anyone when a machine does something that was truly unforeseeable? If HRW/IHRC really wants to impose criminal liability for unforeseeable events, then their issue is not with autonomous weapons, it is with the fundamental precepts of criminal law in rule of law countries.36

Rather, a commander must have a reasonable understanding of the autonomous weapon system and how it will work before deploying it in a particular situation. In addition, as Peter Margulies has said, commanders should exercise what he calls

34 Id.
35 MIND THE GAP, at 19.
36 In the United States for example, due process forbids vicarious criminal liability for acts which are not reasonably foreseeable. See e.g., United States v. Pinkerton, 328 U.S. 640 (1946).
“dynamic diligence,” which is a regime which “will require frequent, periodic assessment and, where necessary, adjustment of [autonomous] weapons’ systems inputs, outputs, and interface with human service members.”

This approach, Margulies contends, is a practical version of what “meaningful human control would look like if that phrase were deployed to permit autonomy while preserving checks on autonomy’s excesses.”

What degree of knowledge a commander must have about the workings of an autonomous weapons’ system in order to be considered reasonable? As it happens, the experts who convened to study the application of the law of war to a related technology – cyber – made a number of relevant observations:

Commanders or other superiors in the chain of command cannot be expected to have a deep knowledge of cyber operations; to some extent, they are entitled to rely on the knowledge and understanding of their subordinates. Nevertheless, the fact that cyber operations may be technically complicated does not alone relieve commanders or other superiors of the responsibility for exercising control over subordinates. Of course, willful or negligent failure to acquire an understanding of such operations is never a justification for lack of knowledge. As a matter of law, commanders and other superiors are assumed to have the same degree of understanding as a ‘reasonable’ commander at a comparable level of command in a similar operational context. In all cases, the knowledge must be sufficient to allow them to fulfill their legal duty to act reasonably to identify, prevent, or stop the commission of cyber war crimes.

There is utterly no reason not to apply this same reasoning to autonomous weapons. This means that in order for designers, commanders, operators and others involved with autonomous weapons to avoid liability, the devices – like any weapon - must be designed and tested so that their expected actions against life and property could be reasonably anticipated. This is nothing new to law of war practitioners. For example, the interpretation of the ICRC of Article 36 of the Additional Protocol clearly indicates that testing is part of the required review.

---

38 Id., at 26.
39 THE TALLINN MANUAL ON INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE 94 (Michael N. Schmitt ed., 2013) [hereinafter TALLINN MANUAL].
40 See note 15, supra.
process for weapons of every sort.\textsuperscript{41}

The U.S. is quite demanding in this regard in order to ensure that weapons are built and used in the proper manner:

[DoD] policy establishes rigorous standards for system design, testing of hardware and software, and training of personnel on the proper use of autonomous and semi-autonomous systems. Among other things, the policy requires that military commanders use autonomous and semi-autonomous weapon systems in a manner consistent with their design, testing, certification, operator training, and doctrine.\textsuperscript{42}

All of this points to the fact that nations that adhere to the law are going to do “rigorous” testing and examination of autonomous weapons so that they do have a reasonable understanding of how they work, and the foreseeable consequences of their use. Absent such ‘due diligence’, those who use those weapons are liable for the consequences if they perform inconsistent with the law of war.

In truth, it is not complicated to find command accountability for directing the use of any weapon without a reasonable belief that doing so would comply with the law of war. Commanders are, after all, expected to take “all necessary and reasonable measures in their power” to prevent war crimes,\textsuperscript{43} and that implicitly requires a reasonable understanding of the foreseeable consequences - something that can be achieved via testing and other processes. In other words, HWR/IHRC’s belief that there can be no accountability because, in their view, autonomous weapons can act “unforeseeably” is obviously wrong because deploying a weapon that is expected to launch attacks “unforeseeably” is itself a punishable breach of the responsibilities of commanders, operators, and the nations they represent.

This is not to say that weapons do not go awry from time to time, but that is not now, nor ever has been, a crime provided reasonable steps have been taken to avoid such an unintended result. It is a fact of war that weapons do not always operate as intended.\textsuperscript{44} But holding someone criminally accountable when a weapon produces


\textsuperscript{42}DOD MANUAL, ¶6.5.9.4 (citing U.S. DEP’T OF DEFENSE DIRECTIVE 3000.09, Autonomy in Weapons Systems (Nov. 21, 2012)) (emphasis added).


\textsuperscript{44}Ulrike Dauer, German City Evacuated After Unexploded World War II Bomb Is Found,
consequences that were unforeseeable despite a rigorous regime of testing that indicated that the device would perform in compliance with the law is simply unjust.

HRW/IHRC also apparently thinks that “criminal liability would likely apply only in situations where the humans specifically intended to use the robots to violate the law.” That is just not how criminal law works. For example, under U.S.’s Uniform Code of Military Justice (Code)\(^\text{45}\) – which in this respect is much like the criminal law of civilian jurisdictions around the globe – the death of another human being can be criminalized in a number of ways.

To illustrate: under Article 118 of the Code criminal liability can be found where the accused kills under circumstances where he engages in “an act which is inherently dangerous to another and evinces a wanton disregard of human life.”\(^\text{46}\) This is but one way a member of the armed forces could be punished for loosing a lethal autonomous weapon without verifying that its operational parameters would comply with the law of war.

Article 119 (manslaughter) criminalizes behavior wherein the accused “who, without an intent to kill or inflict great bodily harm” nevertheless “unlawfully kills a human being…by culpable negligence.”\(^\text{47}\) Thus, involuntary manslaughter may be established by “a negligent act or omission which, when viewed in the light of human experience, might foreseeably result in the death of another, even though death would not necessarily be a natural and probable consequence of the act or omission.”\(^\text{48}\) Is it really difficult to see how criminal liability could be imposed on


\(^{48}\) Culpable negligence is explained as follow:

Culpable negligence is a degree of carelessness greater than simple negligence. It is a negligent act or omission accompanied by a culpable disregard for the foreseeable consequences to others of that act or omission. Thus, the basis of a charge of involuntary manslaughter may be a negligent act or omission which, when viewed in the light of human experience, might foreseeably result in the death of another, even though death would not necessarily be a natural and probable consequence of the act or omission. Acts which may amount to culpable negligence include negligently conducting target practice so that the bullets go in the direction of an inhabited house within range; pointing a pistol in jest at another and pulling the trigger, believing, but without taking reasonable precautions to ascertain, that it would not be dangerous; and carelessly leaving poisons or dangerous drugs where they may endanger life.
anyone involved in the culpably negligent use of an autonomous weapon?\footnote{49}

Additionally, in a provision somewhat unique to U.S. military law, criminal liability can also be imposed upon an accused who causes death merely through \textit{simple} negligence - even in the absence of any intent to kill or injure.\footnote{50} “Simple” negligence is defined as follows:

Simple negligence is the absence of due care, that is, an act or omission of a person who is under a duty to use due care which exhibits a lack of that degree of care of the safety of others which a reasonably careful person would have exercised under the same or similar circumstances. Simple negligence is a lesser degree of carelessness than culpable negligence.\footnote{51}

As an example of the extent to which criminal liability can be extended, an accused was convicted of negligent homicide merely because he lent “his car to a drunken driver who kills himself in an automobile accident”\footnote{52} Consequently, there can be no dispute that a military member who employs (or allows others to employ) an autonomous weapon without being reasonably sure it could and would be used in a way that complies with the law of war \textit{can} be held accountable. In \textit{United States v. Kick},\footnote{53} the all-civilian Court of Military Appeals, explained why, in the military setting, it was necessary to criminalize behavior which breached the relatively low standard of simple negligence.

There is a special need in the military to make the killing of another as a result of simple negligence a criminal act. This is

\begin{quote}
MANUAL FOR COURTS-MARTIAL, UNITED STATES pt.IV, ¶ 44c (2) 9a) (1) (2012) [hereinafter MCM], \textit{available} at \url{http://armypubs.army.mil/epubs/pdf/mcm.pdf}.
\end{quote}

\footnote{49} I disagree with Rebecca Crootof who argues:

Autonomous weapon systems may commit a serious violation of international humanitarian law without anyone acting intentionally or recklessly. Absent such willful action, no one can--or should--be held criminally liable.

\begin{quote}
Rebecca Crootof, \textit{War Torts}, Feb. 2, 2016, Social Studies Research Network, \textit{available at file:///C:/Users/cjd29/Downloads/SSRN-id2657680.pdf}. Among other things, as this article illustrates, U.S. military law has long provided for criminal liability for simple negligence, a markedly lower standard of culpability than recklessness.
\end{quote}

\footnote{50} MCM, \textit{supra} note 48, at pt. IV, ¶85 c (1).

\footnote{51} \textit{Id}. at pt. IV, ¶85 c (2).


because of the extensive use, handling and operation in the course of official duties of such dangerous instruments as weapons, explosives, aircraft, vehicles, and the like. The danger to others from careless acts is so great that society demands protection.\textsuperscript{54}

This illustrates how existing US military law anticipates and recognizes the dangerous potentialities of weaponry, and imposes accountability even where there is the total absence of the “intentiality” HRW/IHRC wrongly thinks must be present to impose criminal liability.\textsuperscript{55} Indeed, this is just a sampling of the myriad of ways that - contrary to what \textit{Mind the Gap} implies - any competent prosecutor could successfully pursue accountably when a fully autonomous weapon is employed.

\begin{center}
\textbf{CIVIL ACCOUNTABILITY}
\end{center}

HRW/IHRC section on civil accountability suffers from a number of conceptual and technical defects. In it is claimed that individual civil damages by victims of an illicit use of an autonomous weapon could not “fill the gap” they perceive to exist in the criminal law. Their discussion mainly centers on the complexity of U.S. tort liability litigation generally, rather than anything to do with weapons’ law or the law of war.

Ironically, in the civil arena, Nevada has passed legislation imposing not just criminal liability in ‘driverless’ car situations, but also civil liability.\textsuperscript{56} And while in theory driverless cars may not be technically ‘fully’ autonomous, as a \textit{practical} matter, they increasingly are so, \textit{de facto}, because a driver’s capability to intervene atrophies over time to the point of ineffectiveness.\textsuperscript{57} Be that as it may, the issue in Nevada seems to be not, as \textit{Mind-the-Gap argument} might suppose, too little potential liability, but too much.\textsuperscript{58}

It is a mistake to underestimate the energy and creativeness of the American plaintiff’s bar. The recent \textit{Arab Bank} case demonstrates that litigants are increasingly figuring out ways to successfully obtain civil judgments for

\begin{flushleft}
\textsuperscript{54} \textit{Id}, at 84.
\textsuperscript{55} \textit{MIND THE GAP}, at 18 (“independent intentionality…must accompany the commission of criminal acts to establish criminal liability”).
\textsuperscript{57} See Frank Douma and Sarah Aue Palodichuk, \textit{Criminal Liability Issues Created by Autonomous Vehicles} 52 SANTA CLARA L. REV. 1157, 1162 (2012).
\textsuperscript{58} Colonna, \textit{supra} note 56
\end{flushleft}
international atrocities and other crimes. There is, however, a dark side to civil litigation is such cases. One scholar wrote in the Georgetown Public Policy Review recently that such civil litigation may benefit individuals, but “the overall effect can be damaging to relationships with key regional partners and weaken the United States’ capacity to investigate terror financing.”

In any event, even if there were an absence of civil liability in the case of a harm caused during an international armed conflict, that is a broader issue than autonomous weapons or, for that matter, any weapon. In fact, internationally speaking, civil liability for even mass torts is extremely problematic. The absence of civil liability of the sort HRW/IHRC seems to think is needed is hardly a reason to ban autonomous weapons.

More specifically as to the law of war, there is virtually no empirical evidence – and Mind the Gap offers none - that civil liability plays any significant role in the shaping of belligerent behavior in the conduct of armed conflict, especially with respect to deterrence. As U.S. District Court Judge Jose A. Cabranes wrote recently in Foreign Affairs, “few evildoers are deterred by the distant threat of monetary damages in civil litigation.”

Of course, it is basic international law that a state which is “responsible for an internationally wrongful act is under an obligation to make full reparation for the injury caused by that act” – and that this principle can apply to law of war violations. The adjudication of such claims is not, however, necessarily bound by the civil law procedures that HRW/IHRC supposes would hamstring such resolutions. Indeed, state parties can settle the claims by whatever procedure they deem appropriate and find mutually acceptable.

That said, it is profoundly misleading to suggest that international law

62 Jose A. Cabranes, Withholding Judgment: Why U.S. Courts Shouldn’t Make Foreign Policy, FOREIGN AFFAIRS, Sept/Oct 2015 125,
63 DOD MANUAL, ¶18.16.
calls for the individual compensation for war crimes. The DoD Manual correctly points out that:

The responsibility of States for violations of the law of war committed by their agents is owed to other States. The fact that such responsibility is owed to other States reflects the predominately inter-State nature of international obligations. *Customary international law and the 1949 Geneva Conventions do not provide a private right for individuals to claim compensation* directly from a State; rather, such claims are made by other States.64

The point is that international relations, not to mention the legitimacy of the law of war, has never depended upon the ability of courts to provide individual compensation, even in the absence of criminal liability. In short, the presence or absence of civil liability is not – and should not be - a necessary condition as to the legitimacy of autonomous weapons.

**CONCLUSION**

It is certainly legitimate – and desirable – to raise questions about autonomous weapons. There are very clearly real complexities associated with the emergence of these devices and their potential uses in warfare.65 However, the notion that there is something intrinsic about them that bars accountability is simply untrue. Manufacturing faux “legal” issues does not advance the dialogue; indeed, as Professor Schmitt suggests, they “muddle” the issues.66

It appears that autonomous weapons’ opponents are grasping at almost any theory to justify a total ban on the technology. Historically, such an approach has proven problematic because as at best, such bans are put in place based on a technological understanding at a specific moment in time. It is quite possible that technology could evolve over time to the point where the ban may actually operate to bar the development or deployment of systems that could operate to save lives of combatants and civilians alike. This is why I have advocated that focus ought to be place not on a particular technology, but rather on strict adherence to the law of war as to its use.67

---

64 Id., at ¶18.16.4
65 The Center for New American Security has established the “Ethical Autonomy Project” to examine these issues, [http://www.cnas.org/ethicalautonomy](http://www.cnas.org/ethicalautonomy) (last visited Mar. 4, 2016).
66 Schmitt, supra, note 32.
67 Charles J. Dunlap, Jr., *To Ban New Weapons or Regulate Their Use?* JUST SECURITY (Apr. 3, 2015, 12:24 PM), [http://justsecurity.org/21766/guest-post-ban-weapons-regulate-use/](http://justsecurity.org/21766/guest-post-ban-weapons-regulate-use/). See also,
The fact of the matter is that these weapons are here to stay. As one former Army officer recognizes:

The technology is already here, and advances in AI [artificial intelligence] in general will create an environment where the continuous development of defensive capabilities will be mandatory. We can’t un-invent deep learning, image recognition algorithms, and supercomputers — despite the FLI’s sincere but misguided attempt to stop advancements in autonomous weapon system development. 68

Among some NGOs and others in the international community there seems to be an instinctive hostility to any technological advance in warfighting, despite the paucity of evidence that increased lethality of weapons necessarily causes more civilians to die violently. In fact, Ian Morris has argued persuasively that in the long run “wars make us safer and richer,” because they force the societal organization and sophistication that ultimately functions to suppress human violence. 69

It is worth noting that the march of time toward a safer society that Morris examines parallels the increasing technological nature of weaponry. Moreover, Martin Van Creveld points out that even with respect to a weapon with such horrific potential as a nuclear bomb, the existence of these lethal instruments has resulted in the disappearance of the most deadly form of conflict: major interstate war. 70 High-tech is not necessarily to be feared as inextricably endangering civilians.

That being the case, rather than searching for reasons to ban sophisticated


68 Sam Wallace, The proposed ban on offensive autonomous weapons is unrealistic and dangerous, KURZWEIL ACCELERATING INTELLIGENCE (Aug. 5, 2015), http://www.kurzweilai.net/the-proposed-ban-on-offensive-autonomous-weapons-is-unrealistic-and-dangerous


70 Martin van Creveld, These nuclear weapons are preventing a war, THE TELEGRAPH, (May 26, 2002, 12:01AM BST), http://www.telegraph.co.uk/comment/personal-view/3577065/These-nuclear-weapons-are-preventing-a-war.html.
weapons, we ought to work to find sensible regulations for them, ever conscious of the grim reality that even in the modern era, some of the worst atrocities have been carried out using not some piece of high-tech weaponry, but the most primitive of implements.\textsuperscript{71} It equally true that in the main, the militaries – such as those of the U.S. and its allies - that have actually tried to suppress man’s inhumanity to man, have come to rely to a great degree upon high-tech weaponry to apply force as precisely as possible in order to spar civilians.

As John Stuart Mill observes, “as long as justice and injustice have not terminated their ever-renewing fight for ascendancy in the affairs of mankind, human beings must be willing, when need is, to do battle for the one against the other.”\textsuperscript{72} We ought not be working to ban the law-abiding nations from seeking to have robots, rather than their young people serving in uniform, doing some of that fighting for justice for us.

* * *

\textsuperscript{71} See \textit{e.g.}, Jean Hatzfeld, \textit{Machete Season: The Killers in Rwanda Speak} (2206).
ADVANCING THE DEBATE ON KILLER ROBOTS:
12 KEY ARGUMENTS FOR A PREEMPTIVE BAN ON FULLY AUTONOMOUS WEAPONS

May 2014

The development of fully autonomous weapons, also known as “killer robots,” and the proposal to ban them preemptively have sparked impassioned debate at the international and national levels. Experts—including lawyers, ethicists, military specialists, human rights advocates, and scientists—have argued about the legality and desirability of the weapons in official diplomatic meetings, at conferences around the world, in academic journals, and on the Internet. In May 2014, states parties to the Convention on Conventional Weapons (CCW) will convene in Geneva for the latest foray into the issue, a four-day experts’ meeting on what CCW states call “lethal autonomous weapons systems.” This paper seeks to advance the discussions about fully autonomous weapons by elaborating on the call for a ban and addressing head on the main arguments against such a ban.

Fully autonomous weapons, once deployed, would be able to select and fire on targets without meaningful human involvement. Although they do not yet exist, the development of precursors and military planning documents indicate that technology is moving rapidly in that direction.

Human Rights Watch and Harvard Law School’s International Human Rights Clinic (IHRC) have contributed to the global discussion with a series of papers arguing for a ban on the development, production, and use of fully autonomous weapons. In November 2012, we released Losing Humanity: The Case against Killer Robots, the first major civil society report on the topic. We subsequently expanded our arguments in other publications, including an analysis of the US Department of Defense’s directive on autonomous weapons, a Q&A document on fully autonomous weapons, a memorandum on the need for new law to ban these weapons, and a report on the human rights implications of the weapons.1

Human Rights Watch and IHRC are calling on governments to:

- Work toward an international instrument prohibiting the development, production, and use of the fully autonomous weapons.
- Develop national policies on the issue, which encompass national moratoria on the development, production, and use of the fully autonomous weapons.
- Agree in November 2014 to expand CCW discussions in a more formal group of governmental experts next year, with an eye ultimately to negotiating a protocol on the weapons.

Our reports on fully autonomous weapons are part of a growing movement against the weapons. The Campaign to Stop Killer Robots, an international coalition of nongovernmental organizations (NGOs) coordinated by Human Rights Watch, has led civil society’s efforts to ban the weapons. It currently has 51 member organizations from 24 countries. Other experts, including Christof Heyns, the UN special rapporteur on extrajudicial, summary or arbitrary executions, have also expressed serious concerns about the legal and moral implications of these weapons.

At the same time, critics of the campaign’s position have defended the proposed technology and challenged the call for a preemptive prohibition. This paper responds directly to those critics by examining and rebutting 12 of their claims. In so doing, it seeks to add depth and nuance to the case against these weapons.

The paper is divided into 12 sections, each providing a response to a particular claim or argument that critics of a preemptive ban have made.

---

Claim #1: A new treaty is unnecessary because existing international humanitarian law is adequate.

**Response:** A new treaty would help clarify existing international humanitarian law and would address development and production of fully autonomous weapons in addition to their use.

**Analysis:** Critics of a new treaty on fully autonomous weapons often assert that existing international humanitarian law is sufficient to deal with the dangers posed by the weapons. They argue that any problematic use of fully autonomous weapons would already be unlawful because it would violate current international humanitarian law. According to two authors, “The question for the legal community [would be] whether autonomous weapon systems comply with the legal norms that States have put in place.”2 Recognizing that the weapons raise new concerns, another author notes that “as cases and mistakes arise, the lawyers and injured parties will have to creatively navigate the network of legal mechanisms [available in international law].”3 Yet he too concludes that a new legal instrument would be unnecessary.

Existing international humanitarian law, however, was not intended to and cannot adequately address the issues raised by fully autonomous weapons. International humanitarian law should be supplemented with new law designed to deal with the unique challenges of this revolutionary type of weapon.

A new international treaty would clarify states’ obligations and make explicit the requirements for compliance. It would minimize questions about legality by standardizing rules across countries and reducing the need for case-by-case determinations. Greater legal clarity would lead to more effective enforcement because countries would better understand the rules. A ban convention would make the illegality of fully autonomous weapons clear even for countries that do not conduct legal reviews of new or modified weapons. Finally, many states that do not join the new treaty would still be apt to abide by its ban because of the stigma associated with the weapons.

---

A treaty dedicated to fully autonomous weapons could also address aspects of proliferation not covered under traditional international humanitarian law, which focuses on the use of weapons in war. In particular, such an instrument could prohibit development and production. Eliminating these activities would prevent the spread of fully autonomous weapons, including to states or non-state actors with little regard for international humanitarian law or limited ability to enforce compliance. In addition, it would help avert an arms race by stopping development before it went too far.

Finally, new law could address concerns about an accountability gap under existing international humanitarian law (see more detailed discussion of accountability under claim #6 below). A treaty could establish a ban under any circumstances and specify that anyone violating that rule would be held responsible for the fully autonomous weapon’s actions.

While international humanitarian law already sets limits on problematic weapons and their use, responsible governments have in the past found it necessary to supplement existing legal frameworks for weapons that by their nature pose significant humanitarian threats. Treaties dedicated to specific weapons types exist for cluster munitions, antipersonnel mines, blinding lasers, chemical weapons, and biological weapons. Fully autonomous weapons have the potential to raise a comparable or even higher level of humanitarian concern and thus should be the subject of similar supplementary international law.

**Claim #2:** Continued developments in artificial intelligence might make it possible for fully autonomous weapons to comply with the principles of distinction and proportionality, at least in certain circumstances.

**Response:** It is likely that fully autonomous weapons would never be capable of reliably complying with the principles of distinction and proportionality.

**Analysis:** Critics argue that advocates of a ban often “fail to take account of likely developments in autonomous weapon systems technology.” According to these critics, not only has military technology “advanced well beyond simply being able to spot an individual or object,” but improvements in artificial intelligence will probably also

---

Thus, while recognizing the existence of “outstanding issues” and “daunting problems,” critics are content with the belief that solutions are “theoretically achievable.” Human Rights Watch and IHRC do not share this optimism and see problems in assuming that such weapons can ever conform to the international humanitarian law requirements of distinction and proportionality.

**Distinction**

Fully autonomous weapons would face great, if not insurmountable, difficulties in reliably distinguishing between lawful and unlawful targets as required by international humanitarian law. Although progress is likely in deep sensory and processing capabilities for these weapons, replicating the elements of human judgment is far more difficult to achieve. The weapons would lack human qualities that facilitate making such determinations, particularly on contemporary battlefields where combatants often seek to conceal their identities. Distinguishing an active combatant from a civilian or injured or surrendering soldier requires more than deep sensory and processing capabilities. It also depends on the qualitative ability to gauge human intention, which involves interpreting subtle, context-dependent clues, such as tone of voice, facial expressions, or body language. Humans possess the unique capacity to identify with other human beings and are thus better equipped to understand the nuances of unforeseen behavior in ways in which machines—which must be programmed in advance—simply cannot.

**Proportionality**

The obstacles presented by the principle of distinction are compounded when it comes to proportionality, which requires the delicate balancing of two factors: expected civilian harm and anticipated military advantage. This evaluation takes place not only in anticipation of an overall military battle plan but also during actual military operations, when decisions must be made about the course or cessation of any particular attack. When it comes to expected civilian harm, one critic has concluded that there “is no question that autonomous weapon systems could be programmed ... to determine the likelihood of harm to civilians in the target area.” Similarly, while

---


8 Ibid., p. 20.
acknowledging that “it is unlikely in the near future that ... ‘machines’ will be programmable to perform robust assessments of a strike’s likely military advantage,” the same critic has written that “military advantage algorithms could in theory be programmed into autonomous weapon systems.”

There are a number of reasons seriously to doubt each of these conclusions. As already discussed, it is highly questionable whether a fully autonomous weapon could ever reliably distinguish legitimate from illegitimate targets. This doubt is enhanced where it is not only the legitimacy of the target that is in question, but also the expected civilian harm—a calculation that requires determining the status of and impact on all the possible entities and objects surrounding the target that might be affected by an attack.

When it comes to predicting anticipated military advantage, even critics admit that “doing so will be challenging [for a machine] because military advantage determinations are always contextual.” More specifically, because military advantage must be determined on a “case-by-case” basis, it is unclear how a programmer could account, in advance, for the infinite variety of unexpected contingencies that may arise in a deployment.

Even if the elements of military advantage and expected civilian harm could be adequately quantified by a fully autonomous weapon system, such a system would be unlikely to be able qualitatively to balance them. The generally accepted standard for assessing proportionality is whether a “reasonable military commander” would have launched a particular attack.

In weighing the proportionality of an attack by a fully autonomous weapon, the appropriate question would be whether the weapon system made a reasonable targeting determination at the time of its strike. Some have suggested the pertinent question with fully autonomous weapons is whether a human commander acted reasonably in deploying it ahead of the strike. The proportionality of any particular

---

9 Ibid. (emphasis added).
11 For a discussion of the case-by-case nature of proportionality, see ibid., p. 256 (asserting that “the military advantage element of the proportionality rule generally necessitates case-by-case determinations”).
13 See Schmitt and Thurnher, ““Out of the Loop,”” Harvard National Security Journal, p. 280 (“Human operators, not machines or software, will ... be making the subjective determinations required under the law of armed conflict, such as those involved in proportionality or precautions in attack calculations. Although the subjective decisions may
attack cannot be ensured at the time of deployment, however, and the decision to deploy is not an equivalent determination to the decision to attack. A commander considering whether to deploy a fully autonomous weapon would need to rely on the programmer’s and manufacturer’s predictions of how the weapon would perform in potentially shifting or unforeseeable conditions. No matter how much care was taken, a programmer or manufacturer would be unlikely to anticipate all conditions that would affect the machine’s performance, which would exacerbate the challenge the commander would face in determining proportionality. In addition, advance programming and testing can never replace human control in the course of operations, and the human judgment to deploy a weapon is not the same as that to mount, continue, alter, or terminate an attack. Yet at the moment of making a determination to attack, a fully autonomous weapon would neither be under the control of a human being exercising his or her own judgment nor able to exercise genuine human judgment itself.

It would be difficult to create machines that could meet the reasonable military commander standard and be expected to act “reasonably” when making determinations to attack in unforeseen or changeable circumstances. According to the Max Planck Encyclopedia of International Law, “[t]he concept of reasonableness exhibits an important link with human reason,” and it is “generally perceived as opening the door to several ethical or moral, rather than legal, considerations.”14 Two critics of the proposed ban treaty have noted that “[p]roportionality ... is partly a technical issue of designing systems capable of measuring predicted civilian harm, but also partly an ethical issue of attaching weights to the variables at stake.”15 Many people would object to the idea that machines could or should be making ethical or moral determinations, and yet this is precisely what the reasonable military commander standard requires. Moreover, reasonableness eludes “objective definition” and depends on the situation.16 If humans cannot know in advance what would be reasonable in every given situation (because making determinations is context-

---

16 Corten, “Reasonableness in International Law,” Max Planck Encyclopedia of Public International Law, para. 1.
specific), it is unrealistic to expect programmers to program machines to act reasonably in inherently unforeseeable situations.

While proportionality analyses allow for a “fairly broad margin of judgment,” the sort of judgment required in deciding how to weigh civilian harm and military advantage in unanticipated situations would be difficult to replicate in machines. As Christof Heyns, the UN special rapporteur on extrajudicial, summary or arbitrary executions, has explained, assessing proportionality requires “distinctively human judgement.” According to the International Committee of the Red Cross (ICRC), judgments about whether a particular attack is proportionate “must above all be a question of common sense and good faith,” characteristics that many would agree machines cannot possess, however thorough their programming.

Claim #3: Fully autonomous weapons should not be treated as per se unlawful because they could be used lawfully in some circumstances.

Response: Narrowly constructed hypothetical cases in which fully autonomous weapons could lawfully be used should not be employed to legitimize the weapons or stand in the way of a ban because the cases do not alter the underlying concerns about the use of such weapons.

Analysis: Critics argue that fully autonomous weapons would not be unlawful per se because there are some potential uses, no matter how limited or unlikely, where they would be both militarily valuable and capable of conforming to the requirements of international humanitarian law. One critic, for example, notes that “[n]ot every battlespace contains civilians.” Other critics maintain fully autonomous weapons

---

19 International Committee of the Red Cross, Commentary of 1987 on Article 57 of Protocol I, para. 2208 (emphasis added).
could be used lawfully under “limited circumstances,” such as in attacks on “nuclear-tipped mobile missile launchers, where millions of lives were at stake.”

One can almost always construct a hypothetical situation where use of a widely condemned weapon could arguably comply with the general rules of international humanitarian law. Before adoption of the 2008 Convention on Cluster Munitions, proponents of cluster munitions often maintained that the weapons could be lawfully launched on a military target alone in an otherwise unpopulated desert. It is extremely difficult, however, to restrict use of weapons to narrowly constructed scenarios, as exemplified by the widespread use of cluster munitions in populated areas. Such theoretical possibilities should not be used to legitimize weapons, including fully autonomous ones, that pose significant humanitarian risks when used in less exceptional situations. The small chance of lawful use in limited circumstances should also not stand in the way of an international prohibition.

Claim #4: Due to the legal requirement to take precautions in attacks, militaries would only use fully autonomous weapons if they were the most humanitarian option among equally effective weapons.

Response: Users of fully autonomous weapons would likely disregard the obligation to take precautions in some circumstances.

Analysis: Some critics object to a ban because they argue that international humanitarian law’s rule on taking precautions in attacks would limit the use of fully autonomous weapons to situations in which they were the most humanitarian alternative. This rule requires militaries to “take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event to minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects.”

Under this provision, critics contend, fully autonomous weapons would only be used if they were the option that would have the least civilian impact among weapons that would achieve the “desired military objective.” One author explains that “the only

---


situation in which an autonomous weapon system can lawfully be employed is when its use will realize military objectives that cannot be attained by other available systems that would cause less collateral damage.\textsuperscript{24}

The critics’ position, however, depends on fully autonomous weapons being used only in compliance with international humanitarian law. Some parties might have little respect for this body of law, while others who generally comply with it might be tempted to use their fully autonomous weapons in certain dire circumstances (see further discussion under claim \#10). The critics ignore these scenarios and the magnitude of the consequences if militaries used the weapons when they were not the most humanitarian option.

A ban would strive to prevent such situations by preempting the creation and proliferation of fully autonomous weapons as well as their use. It would also strongly stigmatize the weapons, putting political pressure on all parties not to use them.

**Claim \#5: Fully autonomous weapons would not be negatively influenced by human emotions when making determinations to use lethal force.**

**Response:** Fully autonomous weapons would lack emotions, including compassion and a resistance to killing, that can protect civilians and soldiers.

**Analysis:** Critics argue that fully autonomous weapons’ lack of human emotions could have military and humanitarian benefits. The weapons would be immune from factors, such as fear, anger, pain, and hunger, that can cloud judgment, distract humans from their military missions, or lead to attacks on civilians.\textsuperscript{25} While such observations have some merit, the role in warfare of other human emotions can in fact advance humanitarian protection in armed conflict.

Humans possess empathy and compassion and are generally reluctant to take the life of another human. A retired US Army Ranger who has done extensive research on killing during war has found that “there is within man an intense resistance to killing

\textsuperscript{24} Ibid.

their fellow man. A resistance so strong that, in many circumstances, soldiers on the battlefield will die before they can overcome it.”

Another author writes,

One of the greatest restraints for the cruelty in war has always been the natural inhibition of humans not to kill or hurt fellow human beings. The natural inhibition is, in fact, so strong that most people would rather die than kill somebody.

Studies of soldiers’ conduct in past conflicts provide evidence to support these conclusions. Human emotions are thus an important inhibitor to killing people unlawfully or needlessly.

Studies have focused largely on troops’ reluctance to kill enemy combatants, but it is reasonable to assume that soldiers feel even greater reluctance to kill the bystanders of war, including civilians or those hors de combat, such as surrendering or wounded soldiers. Fully autonomous weapons, unlike humans, would lack such emotional and moral inhibitions, which, while not required by international law, help protect individuals who are not lawful targets in an armed conflict. One expert writes, “Taking away the inhibition to kill by using robots for the job could weaken the most powerful psychological and ethical restraint in war. War would be inhumanely efficient and would no longer be constrained by the natural urge of soldiers not to kill.”

Due to their lack of emotion, fully autonomous weapons could be the perfect tools for leaders who seek to oppress their own people or to attack civilians in enemy countries. Even the most hardened troops can eventually turn on their leader if ordered to fire on their own people or to commit war crimes. An abusive leader who can resort to fully

28 For example, based on interviews with thousands of US soldiers in World War II, US Army Brig. Gen. S.L.A. Marshall found that usually only 15 to 20 percent of troops would fire at the enemy. These numbers were due to an innate hesitancy to kill, not to fear or cowardice, because “[t]hose who would not fire did not run or hide (and in many cases they were willing to risk great danger to rescue comrades, get ammunition, or run messages).” S.L.A. Marshall, Men against Fire: The Problem of Battle Command in Future War (New York: William Morrow & Company, 1947), p. 54; Grossman, On Killing, p. 4. Other researchers have documented how troops avoided killing by repeatedly loading their guns without firing or by shooting over the enemies’ heads. For discussion of troops in US Civil War repeatedly loading their rifles, see Grossman, On Killing, pp. 18-28. For discussion of Ardant du Picq’s study on nineteenth-century French troops firing in the air, see Grossman, On Killing, pp 9-10. See also Grossman, On Killing, pp. 16-17 (discussing a 1986 study by British Defense Operational Analysis Establishment of 100 “nineteenth- and twentieth-century battles and test trials”).
29 Krishnan, Killer Robots, p. 130.
autonomous weapons would be free of the fear that armed forces would resist being deployed against certain targets.

For all the reasons outlined above, rather than being understood as irrational influences and obstacles to reason, emotions should instead be viewed as central to restraint in war.

Claim #6: Existing international humanitarian law can adequately address accountability concerns arising out of the use of fully autonomous weapons.

Response: Insurmountable legal and practical obstacles would likely interfere with holding someone accountable for unforeseeable, unlawful acts committed by a fully autonomous weapon.

Analysis: Critics argue that the “mere fact that a human might not be in control of a particular engagement does not mean that no human is responsible for the actions of the autonomous weapon system.” According to these critics, “[a] human must decide how to program the system and when to launch it.” Thus, a programmer or manufacturer could be held accountable for intentionally creating a robot that would commit war crimes, and the person deploying the robot “would be accountable for those war crimes if he or she knew or should have known that the autonomous weapon system had been so programmed and did nothing to stop its use.”

Certainly a commander could be held responsible for intentionally using a fully autonomous weapon that was clearly unsuited for the environment in which it was deployed. Furthermore, a programmer or manufacturer could be held liable for intentionally producing a fully autonomous weapon that would commit war crimes. These scenarios, however, fail to capture perhaps more likely situations in which the commander, programmer, or manufacturer did not know a robot would commit an

31 Ibid.
32 Ibid.
33 Ibid. (“Hopefully, autonomous weapon systems will never be programmed to commit war crimes. Much more likely would be a case in which a system that has not been so programmed is nevertheless used in a manner that constitutes such crimes. For example, the operator of an autonomous weapon system that cannot distinguish civilians from combatants who employs the system in an area where the two are intermingled has committed the war crime of indiscriminate attack.”).
illegal act, but the robot nonetheless unexpectedly did so. In such cases, there would be no human to hold directly responsible for the decision to attack, and indirect liability would be difficult to achieve.

Due to the autonomous nature of their technology, fully autonomous weapons could act unforeseeably in ways that would cause unlawful harm. One type of war crime is the act of willfully “launching an indiscriminate attack affecting the civilian population or civilian objects in the knowledge that such attack will cause excessive loss of life, injury to civilians or damage to civilian objects.” Given the technological limitations discussed under claim #2, a fully autonomous weapon, created and deployed to comply with international humanitarian law, might nonetheless launch an attack whose military advantage was so obviously outweighed by its civilian costs that any reasonable human in its position would have known that the attack was disproportionate. Similarly, a fully autonomous weapon might select and fire upon a civilian target mistaking it for a military target, even when a reasonable human would have known that the object of the attack was civilian.

Significant obstacles to holding anyone accountable would exist for both of these situations. Robots themselves could not be punished for committing war crimes as they lack the capacity to feel pain or other emotions associated with punishment. Command responsibility holds military commanders responsible for subordinates’ actions if they knew or should have known their subordinates committed or were going to commit a crime and failed to prevent the crime or punish the subordinates. In the circumstances described above, the commander could not foresee and thus not prevent the violations in question, and he or she could not punish the robot after the fact.

An alternative option would be to try to hold the programmer or manufacturer civilly liable for the unanticipated acts of a fully autonomous weapon. Tort law offers an approach other than prosecution, but it too would likely fail to ensure accountability. In the United States, for example, defense contractors are generally not found liable for harm caused by their products. Even without a legal gap, there are policy and

34 Protocol I, art. 85 (emphasis added).
36 Under the Federal Torts Claims Act, the government waives its immunity from civil suits in certain situations. The Supreme Court has applied this rule to contractors hired by the government. The waiver, however, is subject to the discretionary function exception and the combatant activities exception, which would block most suits against
practical problems with holding programmers and manufacturers accountable. Such liability could be unfair since even programmers and manufacturers might be unable to foresee the harm their fully autonomous weapons could cause in various situations.\(^{37}\) In addition, civil suits are generally brought by victims and, especially in cases of armed conflict, it is unrealistic to think all victims would have the resources or adequate access to obtain justice. This practical limitation is significant because civil litigation against those who program, manufacture, or use such robots would be a more likely avenue of redress than prosecution.\(^{38}\)

The use of fully autonomous weapons would thus lead to the creation of a potentially insurmountable accountability gap. The lack of criminal or civil consequences would interfere with deterrence. A failure to punish would leave victims and their relatives without the satisfaction that someone paid for the suffering they experienced.

**Claim #7:** The Martens Clause does not restrict the use of fully autonomous weapons.

**Response:** Because existing law does not specifically address the unique issues raised by fully autonomous weapons, the Martens Clause mandates that the “principles of humanity” and “dictates of public conscience” be factored into an analysis of their legality. Both of these standards weigh in favor of a ban on this kind of technology.

**Analysis:** Some critics dismiss the value of the Martens Clause in determining the legality of fully autonomous weapons. As it appears in Additional Protocol I, the Martens Clause mandates that:

---


In cases not covered by this Protocol or by other international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from the dictates of public conscience.39

Critics argue that the Martens Clause “does not act as an overarching principle that must be considered in every case,” but is, rather, merely “a failsafe mechanism meant to address lacunae in the law.”40 They contend that because gaps in the law are rare, “the likelihood that” autonomous weapon systems would violate the Martens Clause but not “applicable treaty and customary law” is therefore “exceptionally low.”41 The lack of specific law on fully autonomous weapons, however, means that the Martens Clause does apply, and the weapons raise serious concerns under the provision.

The key question in determining the relevance of the Martens Clause to fully autonomous weapons is the extent to which such weapons are “covered” by existing treaty law. As the US Military Tribunal at Nuremberg explained, the Martens Clause makes “the usages established among civilized nations, the laws of humanity and the dictates of public conscience into the legal yardstick to be applied if and when the specific provisions of [existing law] do not cover specific cases occurring in warfare.”42 The International Court of Justice asserted that the clause’s “continuing existence and applicability is not to be doubted” and that it has “proved to be an effective means of addressing the rapid evolution of military technology.”43 Fully autonomous weapon systems are rapidly evolving forms of technology that are at best only generally covered by existing law.44

39 Protocol 1, art 1(2).
41 Ibid., p. 276.
44 Some critics argue that international humanitarian law would adequately cover autonomous weapon systems, but the most relevant rules are general ones, such as those of distinction and proportionality discussed above under claim #2. While critics also emphasize the applicability of disarmament treaties on antipersonnel landmines, cluster munitions, and incendiary weapons, these instruments do not provide specific law on fully autonomous weapons. They would only govern the payload of narrow group of such weapons and would not address the challenging issues unique to autonomous systems. To date, there is no specific law dedicated to fully autonomous weapons. For critics’ view, see Schmitt and Thurnher, “‘Out of the Loop,’” Harvard National Security Journal, p. 276.
The plain language of the Martens Clause elevates the “principles of humanity” and the “dictates of public conscience” to independent legal standards against which new forms of military technology should be evaluated. On this basis, any weapon conflicting with either of these standards is therefore arguably unlawful. At a minimum, however, the dictates of public conscience and principles of humanity can “serve as fundamental guidance in the interpretation of international customary or treaty rules.” Thus, “[i]n case of doubt, international rules, in particular rules belonging to humanitarian law, must be construed so as to be consonant with general standards of humanity and the demands of public conscience.” Given the serious doubts about the ability of fully autonomous weapons to conform to the requirements of the law, as discussed under claim #2, the standards of the Martens Clause should at the very least be taken into account when evaluating them.

Fully autonomous weapons raise serious concerns under the principles of humanity and dictates of public conscience. The ICRC has described the principles of humanity as requiring compassion and the ability to protect. The challenges the weapons would face in meeting international humanitarian law suggest they could not adequately protect civilians. As discussed above under claim #5, robots would also lack human emotions, including compassion. Public opinion can play a role in revealing and shaping public conscience, and many people find the prospect of delegating life-and-death decisions to machines shocking and unacceptable. A 2013 national representative survey of 1,000 Americans found that, of those with a view, two-thirds came out against fully autonomous weapons: 68 percent opposed the move toward these weapons (48 percent strongly), while 32 percent favored their

---


development. Interestingly, active duty military personnel were among the strongest objectors—73 percent expressed opposition to fully autonomous weapons. These kinds of reactions indicate that fully autonomous weapons could contravene the Martens Clause.

Concerns about weapons’ compliance with the principles in the Martens Clause have justified new weapons treaties in the past. For example, the Martens Clause heavily influenced the discussions and debates preceding the development of CCW Protocol IV on Blinding Lasers, which bans the transfer and use of laser weapons whose sole or partial purpose is to cause permanent blindness. The Martens Clause was invoked not only by civil society in its reports on the matter, but also by experts participating in a series of ICRC meetings on the subject. They largely agreed that “[blinding lasers] would run counter to the requirements of established custom, humanity, and public conscience.” It appears that a shared visceral reaction against blinding weapons ultimately tipped the scales toward a prohibition, even without consensus that such weapons were unlawful under the core principles of international humanitarian law. The Blinding Lasers Protocol set an international precedent for preemptively banning weapons based at least in part on the Martens Clause. Invoking the clause in the context of fully autonomous weapons would be equally appropriate.


52 According to the ICRC report, “some experts expressed either personal repugnance for lasers or the belief that their countries’ civilian population would find the use of blinding as a method of warfare horrific.” International Committee of the Red Cross, Blinding Weapons: Reports of the Meetings of Experts Convened by the International Committee of the Red Cross on Battlefield Laser Weapons, 1989-1991 (Geneva: International Committee of the Red Cross, 1993), pp. 344-46. Others doubted their ability to field such weapons, notwithstanding possible military utility, because of public opinion. Ibid., p. 341.

53 This visceral reaction is suggested by the comments of the participating experts in the ICRC meetings. Examples include the statement of one participant that he would be unable to introduce blinding weapons in his country “because public opinion would be repulsed at the idea.” Another participant described it as “Indisputable that deliberately blinding on the battlefield would be socially unacceptable.” Ibid., p. 345.

Claim #8: A ban on fully autonomous weapons is premature given the possibility of a technological fix.

Response: These highly problematic weapons should be preemptively banned to prevent serious humanitarian harm before it is too late and to accord with the precautionary principle.

Analysis: Critics contend that a preemptive ban on the development, production, and use of fully autonomous weapons is premature. They argue that:

[research into the possibilities of autonomous machine decision-making, not just in weapons but across many human activities, is only a couple of decades old.... We should not rule out in advance possibilities of positive technological outcomes—including the development of technologies of war that might reduce risks to civilians by making targeting more precise and firing decisions more controlled.]

This position depends in part on one’s faith that technology could address the legal challenges raised by fully autonomous weapons, which, as explained above under claim #2, Human Rights Watch and IHRC question. At the same time, it ignores other problems associated with these weapons that are not related to technology, notably the potential for an arms race, an accountability gap, and moral objections, which are discussed under claims #1, 6, and 9.

Given the host of concerns about fully autonomous weapons, they should be preemptively banned before it becomes too late to change course. It is difficult to stop technology once large-scale investments have been made. The temptation to use technology already developed and incorporated into military arsenals would be great, and many countries would be reluctant to give it up, especially if their competitors were deploying it.

In addition, if ongoing development were permitted, militaries might deploy fully autonomous weapons in complex circumstances before artificial intelligence could handle them. Only after the weapons faced unanticipated situations that they were not programmed to address could the technology be modified to resolve those issues.

During that period, the weapon would be likely to mishandle such situations potentially causing great civilian harm.

The prevalence of humanitarian concerns and the uncertainty regarding technology make it appropriate to invoke the precautionary principle, a principle of international law. The 1992 Rio Declaration states, “Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.”\(^{56}\) While the Rio Declaration applies the precautionary principle to environmental protection, the principle can be adapted to other situations.

Fully autonomous weapons implicate the three essential elements of the precautionary principle—threat of serious or irreversible damage, scientific uncertainty, and the availability of cost-effective measures to prevent harm. The development, production, and use of fully autonomous weapons present a threat to civilians that would be both serious and irreversible, as the technology would revolutionize armed conflict and would be difficult to eliminate once developed and employed. Scientific uncertainty characterizes the debate over these weapons. Defenders argue there is no proof that a technological fix could not solve the problem, but there is an equal lack of proof that a technological fix would work. Finally, while treaty negotiations and implementation would carry costs, these expenses are small compared to the significant harm they might prevent.

There is precedent for a preemptive ban on a class of weapons. As discussed under claim #7 above, in 1995 governments agreed to a ban on blinding lasers before the weapons had started to be deployed out of concerns for the humanitarian harm the weapons would cause.\(^{57}\) While an international, preemptive ban should be the end goal, a first step could be national moratoria on fully autonomous weapons, such as those proposed by Christof Heyns, the UN special rapporteur on extrajudicial, summary or arbitrary executions, in his May 2013 report to the Human Rights Council.\(^{58}\)

-----


Claim #9: Fully autonomous weapons could have military and humanitarian benefits that would be sacrificed by a preemptive ban on such technology.

**Response:** The potential dangers posed by fully autonomous weapons would offset any possible benefits. Furthermore, delegating life-and-death decisions to machines would be deeply problematic from a moral standpoint.

**Analysis:** Critics argue that a preemptive ban on fully autonomous weapons would mean forgoing the potential advantages of such technology. They believe that the weapons could provide military benefits such as faster-than-human data processing and response time and enhanced protection of soldiers.⁵⁹ According to some, greater autonomy could also help prevent civilian harm through better target identification or more accurate estimations of the damage likely to result from a particular attack.⁶⁰

The drawbacks of such technology would outweigh the conceivable benefits, however. For example, the greater speed with which fully autonomous weapons might act, while arguably offering some military advantage, would raise the possibility that armed conflicts could rapidly spiral out of control. In arguing that fully autonomous weapons could become a necessity for states seeking to keep up with their adversaries, two critics write that “[f]uture combat may ... occur at such a high tempo that human operators will simply be unable to keep up. Indeed, advanced weapon systems may well create an environment too complex for humans to direct.”⁶¹ The frightening nature of such a scenario is compounded by the prospect of an arms race in which states increasingly feel the need to rely on fully autonomous weapons. The fact that some have gone so far as to contend that “autonomous weapons may become the norm on

---

⁵⁹ See, for example, Arkin, *Governing Lethal Behavior in Autonomous Robots*, p. 30 (Robots “can integrate more information from more sources far faster before responding with lethal force than a human possibly could in real-time.”); P.W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* (New York: The Penguin Press, 2009), p. 418 (“[M]ost of the focus on military robotics is to use robots as a replacement for human losses.”).

⁶⁰ Anderson and Waxman, “Law and Ethics for Autonomous Weapon Systems,” *Jean Perkins Task Force on National Security and Law*, p. 15 (“It may well be, for instance, that weapons systems with greater and greater levels of automation can—in some battlefield contexts, and perhaps more and more over time—reduce misidentification of military targets, better detect or calculate possible collateral damage, or allow for using smaller quanta of force compared to human decision-making.”). Schmitt and Thurnher, “‘Out of the Loop,’” *Harvard National Security Journal*, p. 3 (“Perhaps even more troubling is the prospect that banning autonomous weapon systems altogether based on speculation as to their future form could forfeit their potential use in a manner that would minimize harm to civilians and civilian objects when compared to non-autonomous weapon systems.”).

the battlefield in a generation” makes it clear just how pressing the issue has become.62

While perceived military advantage may be driving much of the push towards fully autonomous weapons, certain claims regarding such technology’s benefits are overstated. Critics correctly point out that countries have an interest in “fielding systems that enable them to deliver lethal force while minimizing the risk to their own forces,” and saving soldiers’ lives is a laudable goal.63 Existing semi-autonomous systems, such as armed drones, however, provide similar force protection while leaving a human in control of the ultimate firing decision.

The use of fully autonomous weapons also raises the troubling specter that life-and-death decisions will be increasingly delegated to machines that, by their very nature, are without a moral compass. Thus, even if fully autonomous weapons might conceivably someday take actions with “potentially lethal consequences ... better than humans can”64—a development Human Rights Watch and IHRC believe is unlikely—ceding control over lethal force to these weapons would be deeply problematic from a moral perspective. According to Christof Heyns, the UN special rapporteur on extrajudicial, summary or arbitrary executions, “[m]achines lack morality and mortality, and should as a result not have life and death powers over humans.”65 Heyns adds that “[i]t is an underlying assumption of most legal, moral and other codes that when the decision to take life or to subject people to other grave consequences is at stake, the decision-making power should be exercised by humans.”66 The fact that delegating life-and-death decisions to machines is so deeply disturbing across cultures should trump any arguable benefits touted by critics.

---

62 Ibid., p. 239 (describing the conclusions of certain US Department of Defense studies about future norms on the battlefield).
63 Ibid., p. 232.
66 Ibid., pp. 16-17.
Claim #10: Concerns about fully autonomous weapons would be best addressed through a regulatory approach rather than a ban.

Response: A binding, absolute ban on fully autonomous weapons would reduce the chance of misuse of the weapons, be easier to enforce, and enhance the stigma associated with violations.

Analysis: Certain critics object to a categorical ban on fully autonomous weapons because they prefer a regulatory framework that would permit the use of such technology in accordance with certain pre-defined limitations. Such a framework might, for example, restrict the use of fully autonomous weapons to specific types of locations or purposes. These critics suggest that such an approach would more precisely tailor restrictions to the evolving state of fully autonomous weapons technology and thus not be over-inclusive. Regulations could come in the form of a legally binding instrument or a set of gradually developed, informal standards.

An absolute, legally binding ban on fully autonomous weapons, however, would provide several distinct advantages over formal or informal constraints. It would maximize protection for civilians in conflict because it would be more comprehensive than regulation. A ban would also be more effective as it would prohibit the existence of the weapons and be easier to enforce. Moreover, a ban can have a powerful stigmatizing effect, creating a widely recognized new standard and influencing even those that do not join the treaty.

By contrast, once fully autonomous weapons came into being under a regulatory regime, they would be vulnerable to misuse. Even if regulations restricted use of fully autonomous weapons to certain locations or specific purposes, after the weapons entered national arsenals, countries that usually respect international humanitarian law could be tempted to use the weapons in inappropriate ways in the heat of battle or in dire circumstances. Furthermore, the existence of fully autonomous weapons would

---

67 See Armin Krishnan, “Automating War: The Need for Regulation,” Contemporary Security Policy, vol. 30, no. 1 (2009), p. 189 ("The best option of dealing with the possible implications of military robotics is probably not a general ban.... What is proposed in here as a solution is to allow defensive applications of [autonomous weapons], but to put considerable restrictions on offensive types and to ban certain types (self-evolving, self-replicating robots, microrobots) completely.").

68 See, for example, ibid., p. 188.

69 Anderson and Waxman, “Law and Ethics for Autonomous Weapon Systems,” Jean Perkins Task Force on National Security and Law, p. 22 (explaining that “[bly ‘international norms’ here, we do not mean new binding legal rules only—whether treaty rules or customary international law—but instead the gradual fostering of widely-held expectations about legally or ethically appropriate conduct, whether formally binding or not”.

---

Advancing the Debate on Killer Robots 22
leave open the door to their acquisition by repressive regimes or non-state armed
groups that might disregard the restrictions or alter or override any programming
designed to regulate a robot’s behavior. They could use the weapons against their own
people or civilians in other countries with horrific consequences.

Enforcement of regulations on fully autonomous weapons, as on all regulated weapons,
could also be challenging and leave room for error, increasing the potential for harm to
civilians. Instead of knowing that any use of fully autonomous weapons is unlawful,
countries, international organizations, and nongovernmental organizations would have
to monitor the use of the weapons and determine in every case whether use complied
with the regulations. There would probably be debates about enforcement and the
scope of the regulations—for example, what constituted a populated area, where use of
certain weapons might be restricted.

The challenges of effectively controlling the use of fully autonomous weapons through
binding regulations would be compounded if governments adopted a non-binding
option. Those who support best practices advocate “let[ting] other, less formal
processes take the lead to allow genuinely widely shared norms to coalesce in a very
difficult area.” To the extent that a “less formal” approach is a non-binding one, it is
highly unlikely to constrain governments—including those already inclined to violate
the law—in any meaningful way, especially under the pressures of armed conflict. It is
similarly unrealistic to expect governments, as some critics hope, to resist their
“impulses toward secrecy and reticence with respect to military technologies” and
contribute to a normative dialogue about the appropriate use of fully autonomous
weapons technology. If countries rely on transparency and wait until “norms coalesce”
in an admittedly “very difficult area,” such weapons will likely be developed and
deployed, at which point it would probably already be too late to control them.

---

70 Ibid., p. 20.
71 Ibid., p. 25 (referring to the US tendencies toward secrecy).
72 Ibid., p. 20.
Claim #11: Efforts to ban fully autonomous weapons should be abandoned because an international prohibition is unrealistic.

Response: Past disarmament successes, growing support for a ban, and increasing international discussion of the issue suggest that a ban is a viable option for addressing fully autonomous weapons.

Analysis: Some critics argue that an absolute ban on the development, production, and use of fully autonomous weapons is “unrealistic.” They have written that “part of our disagreements are about the practical difficulties that face international legal prohibitions of military technologies (we think such efforts are likely to fail).” These critics fail to acknowledge the parallels with past successful disarmament efforts that had humanitarian benefits and the growing support for preserving meaningful human control over decisions to use lethal force.

Strong precedent exists for banning weapons that raise serious humanitarian concerns. The international community has previously adopted legally binding prohibitions on poison gas, biological weapons, chemical weapons, blinding lasers, and antipersonnel landmines. Most recently, 107 states adopted the 2008 Convention on Cluster Munitions, which comprehensively bans the use, production, transfer, and stockpiling of cluster munitions. Opponents of the landmine and cluster munitions instruments frequently said that a ban treaty would never be possible, but the end results proved that their skepticism was misplaced.

Efforts to address the problems of fully autonomous weapons are following a similar path. April 2013 marked the launch of the Campaign to Stop Killer Robots, which calls for an absolute ban on the development, production, and use of fully autonomous weapons. The campaign resembles earlier civil society coalitions, including the Cluster Munition Coalition and International Campaign to Ban Landmines.

Public support for a ban has bolstered the position of the campaign. In a 2013 public letter, more than 270 roboticists, artificial intelligence experts, and other scientists from 37 countries expressed doubts that adequate technological developments to ensure such weapons would comport with international humanitarian law would be

---

73 Ibid., p. 3.
possible. They wrote, “[G]iven the limitations and unknown future risks of autonomous robot weapons technology, we call for a prohibition on their development and deployment. Decisions about the application of violent force must not be delegated to machines.” Surveys have also revealed public resistance to the prospect of fully autonomous weapons. As mentioned under claim #7 above, a 2013 survey of 1,000 Americans found that two-thirds came out against the weapons, with almost three quarters of active duty military expressing opposition.

Interim measures to address the problem have also received attention. In May 2013, Christof Heyns, the UN special rapporteur on extrajudicial, summary or arbitrary executions, submitted to the Human Rights Council a report that was highly critical of fully autonomous weapons. He argued that the weapons would pose serious legal and ethical dangers and wrote that “[t]hey raise far reaching concerns about the protection of life during war and peace.” Heyns called for national moratoria, which are temporary bans, on fully autonomous weapons and a high level panel to develop an international policy on the issue. While not adopting the language of a moratorium, the US Department of Defense issued a de facto moratorium in a 2012 directive that prohibits for up to 10 years the use of fully autonomous systems to deliver lethal force. Such moratoria can be a first step toward an absolute ban.

Finally, governments have taken up the debate about fully autonomous weapons in key international bodies. The Heyns report generated discussion in the Human Rights Council, where almost 20 nations expressed concern about the weapons and some endorsed the special rapporteur’s call for moratoria. More importantly, the 117 states parties to the CCW agreed to take up the issue in informal discussions in May 2014. While this development does not indicate universal support for a ban, it does reflect a willingness by many major military powers to discuss concerns about the use of fully autonomous weapons in a leading disarmament forum. The CCW process has in the past produced a preemptive ban on blinding lasers and served as an incubator for bans on landmines and cluster munitions. There is certainly much work to be done to

78 Ibid.
achieve a ban, but past precedents and recent developments suggest it cannot be dismissed as unrealistic.

Claim #12: A ban on the development of fully autonomous weapons would impede development of other valuable autonomous technology.

Response: A prohibition would not stifle such advances in autonomous technology because it would not cover non-weaponized, fully autonomous technology or semi-autonomous weapon systems.

Analysis: Some critics worry about the breadth of a ban on development. They express concern that it would represent a prohibition “even on the development of technologies or components of automation that could lead to fully autonomous lethal weapon systems.” These critics fear that the ban would therefore impede the exploration of beneficial autonomous technology, such as self-driving cars.

In fact, the ban would apply to development only of fully autonomous weapons, that is, machines that could select and fire on targets without human intervention. Research and development activities would be banned if they were directed at technology that can be used exclusively for fully autonomous weapons or that is explicitly intended for use in such weapons. A prohibition on the development of fully autonomous weapons would in no way impede development of fully autonomous robotics technology, which can have many positive, non-military applications.

The prohibition would also not encompass development of semi-autonomous weapons such as existing remote-controlled armed drones. Even if the prohibition is a narrow one, as a matter of principle, countries would not, and should not, be permitted to contract specifically for the development of fully autonomous weapon systems.

---

Conclusion

Fully autonomous weapons raise a host of humanitarian concerns that cumulatively outweigh any potential benefits. They would face obstacles to complying with international humanitarian law, have the potential to proliferate and start an arms race, create an accountability gap, undermine non-legal checks on killing, and present the ethical problem of ceding life-and-death decisions to machines. To address these concerns, governments should adopt new international law to supplement existing international humanitarian law. In particular, they should agree to a preemptive ban, rather than regulation. An absolute prohibition on development, production, and use of fully autonomous weapon would be the most effective approach to eliminating the humanitarian threats posed by fully autonomous weapons.
ARTIFICIAL INTELLIGENCE POLICY: A PRIMER AND ROADMAP

Ryan Calo*

The year is 2017 and talk of artificial intelligence is everywhere. People marvel at the capacity of machines to translate any language and master any game.1 Others condemn the use of secret algorithms to sentence criminal defendants2 or recoil at the prospect of machines gunning for blue, pink, and white-collar jobs.3 Some worry aloud that artificial intelligence (“AI”) will be humankind’s “final invention.”4

The attention we pay to AI today is hardly new: looking back twenty, forty, or even a hundred years, one encounters similar hopes and concerns around AI systems and the robots they inhabit. Batya Friedman and Helen Nissenbaum wrote Bias in Computer Systems, a framework for evaluating and responding to machines that discriminate unfairly, in 1996.5 The 1980 New York Times headline “A Robot is After Your Job” could as easily appear in September 2017.6

The field of artificial intelligence itself dates back at least to the 1950s, when John McCarthy and others coined the term one summer at Dartmouth College, and the concepts underlying AI go back generations earlier to the ideas of Charles Babbage, Ada

---

1 Lane Powell and D. Wayne Gittinger Associate Professor, University of Washington School of Law. The author would like to thank a variety of individuals within industry, government, and academia who have shared their thoughts, including Miles Brundage. Anupam Chander, Rebecca Croootof, Oren Etzioni, Ryan Hagemann, Woodrow Hartzog, Alex Kozak, Amanda Levendowski, Mark MacCarthy, Patrick Moore, Julia Powles, Helen Toner, and Eduard Fosch Villaronga. Thanks also to Madeline Lamo for excellent research assistance and to the editors of the UC Davis Law Review for excellent suggestions. The author is co-director of an interdisciplinary lab that receives funding to study emerging technology, including from the Microsoft Corporation, the William and Flora Hewlett Foundation, and the John D. and Catherine T. MacArthur Foundation.

2 See, e.g., Cade Metz, In a Huge Breakthrough, Google’s AI Beats a Top Player at the Game of Go, WIRED (Jan. 27, 2016), https://www.wired.com/2016/01/in-a-huge-breakthrough-googles-ai-beats-a-top-player-at-the-game-of-go/ (reporting how after decades of work, Google’s AI finally beat the top human player in the game of Go, a 2,500 year-old game of strategy and intuition exponentially more complex than chess).

3 See, e.g., MARTIN FORD, RISE OF THE ROBOTS: TECHNOLOGY AND THE THREAT OF A JOBLESS FUTURE (2015) (predicting that machine’s role will evolve from that of the worker’s tool to the worker itself).

4 See generally JAMES BARRAT, OUR FINAL INVENTION: ARTIFICIAL INTELLIGENCE AND THE END OF THE HUMAN ERA (2013) (“Our species is going to mortally struggle with this problem.”).

5 Batya Friedman & Helen Nissenbaum, Bias in Computer Systems, 14 ACM TRANSACTIONS ON INFO. SYST. 330 (1996).

6 Harley Shaiken, A Robot Is After Your Job; New Technology Isn’t a Panacea, N.Y. TIMES, Sept. 3, 1980, at A19. For an excellent timeline of coverage of robots displacing labor, see Louis Anslow, Robots Have Been About to Take All the Jobs for More Than 200 Years, TIMELINE (May 16, 2016), https://timeline.com/robots-have-been-about-to-take-all-the-jobs-for-more-than-200-years-5c9c08a2f41d.
Lovelace, and Alan Turing. Although there have been significant developments and refinements, nearly every technique we use today — including the biologically-inspired neural nets at the core of the practical AI breakthroughs currently making headlines — was developed decades ago by researchers in the United States, Canada, and elsewhere.  

If the terminology, constituent techniques, and hopes and fears around artificial intelligence are not new, what exactly is? At least two differences characterize the present climate. First, as is widely remarked, a vast increase in computational power and access to training data has led to practical breakthroughs in machine learning, a singularly important branch of AI. These breakthroughs underpin recent successes across a variety of applied domains, from diagnosing precancerous moles to driving a vehicle, and dramatize the potential of AI for both good and ill.

Second, policymakers are finally paying close attention. In 1960, when John F. Kennedy was elected, there were calls for him to hold a conference around robots and labor. He declined. Later there were calls to form a Federal Automation Commission. None was formed. A search revealed no hearings on artificial intelligence in the House or Senate until, within months of one another in 2016, the House Energy and Commerce Committee held a hearing on Advanced Robotics (robots with AI) and the Senate Joint Economic Committee held the “first ever hearing focused solely on artificial intelligence.” That same year, the Obama White House held several workshops on AI and published three official reports detailing its findings. Formal policymaking around AI abroad is, if anything, more advanced: the governments of Japan and the European Union have proposed or formed official commissions around robots and AI in recent years.

---

8 See id. at 50-51.; Will Knight, Facebook Heads to Canada for the Next Big AI Breakthrough, Technology Review (Sept. 15, 2017), online at https://www.technologyreview.com/s/608858/facebook-heads-to-canada-for-the-next-big-ai-breakthrough/ (discussing leading figures and breakthroughs with connections to Canada).
10 See Anslow, supra note 6.
11 He did, however, give a speech on the necessity of “effective and vigorous government leadership” to help solve the “problems of automation.” John F. Kennedy, U.S. President, Labor: Meeting the Problems of Automation (June 7, 1960).
12 See Anslow, supra note 6.
14 E.g., Nat’l Sci. & Tech. Council, supra note 9, at 12.
This essay, prepared in connection with the UC Davis Law Review’s 50th Anniversary symposium Future-Proofing Law: from rDNA to Robots, is my attempt at introducing the AI policy debate to recent audiences, as well as offering a conceptual organization for existing participants. The essay is designed to help policymakers, investors, scholars, and students understand the contemporary policy environment around artificial intelligence and the key challenges it presents. These include:

- Justice and equity
- Use of force
- Safety and certification
- Privacy and power; and
- Taxation and displacement of labor.

In addition to these topics, the essay will touch briefly on a selection of broader systemic questions:

- Institutional configuration and expertise
- Investment and procurement
- Removing hurdles to accountability; and
- Correcting flawed mental models of AI.

In each instance, the essay endeavors to give sufficient detail to describe the challenge without prejudging the policy outcome. This essay is meant to be a roadmap, not the road itself. Its primary goal is to point the new entrant toward a wider debate and equip her with the context for further exploration and research.

I am a law professor with no formal training in AI. But my longstanding engagement with AI has provided me with a front row seat to many of the recent efforts to assess and channel the impact of AI on society.16 I am familiar with the burgeoning literature and commentary on this topic and have reached out to individuals in the field to get their sense of what is important. That said, I certainly would not suggest that the inventory of policy questions I identify here is somehow a matter of consensus. I do not speak for the AI policy community as a whole. Rather, the views that follow are idiosyncratic and reflect, in the end, one scholar’s interpretation of a complex landscape.17

---

16 For example, I hosted the first White House workshop on artificial intelligence policy, participated as an expert in the inaugural panel of the Stanford AI 100 study, organized AI workshops for the National Science Foundation, the Department of Homeland Security, and the National Academy of Sciences, advised AI Now and FAT*, and co-founded the We Robot conference.

17 Earlier AI pioneer Herbert Simon argues that it is the duty of people who study a new technology to offer their interpretations regarding its likely effects on society. HERBERT SIMON, THE SHAPE OF AUTOMATION FOR MEN AND MANAGEMENT, vii (1965). But: “Such interpretations should be, of course, the beginning and not the end of public discussion.” Id. I vehemently agree. For another interpretation, focusing on careers in AI policy, see Miles Brundage, GUIDE TO WORKING IN AI POLICY AND STRATEGY, 80,000 HOURS (2017), https://80000hours.org/articles/ai-policy-guide/.
The remainder of the essay proceeds as follows. Part I offers a short background on artificial intelligence and defends the terminology of policy over comparable terms such as ethics and governance. Part II lays out the key policy concerns of AI as of this writing. Part III addresses the oddly tenacious and prevalent fear that AI poses an existential threat to humanity — a concern that, if true, would seem to dwarf all other policy concerns. A final section concludes.

I. Background

What is AI?

There is no straightforward, consensus definition of artificial intelligence. AI is best understood as a set of techniques aimed at approximating some aspect of human or animal cognition using machines. Early theorists conceived of symbolic systems — the organization of abstract symbols using logical rules — as the most fruitful path toward computers that can “think.” But the approach of building a reasoning machine upon which to scaffold all other cognitive tasks, as originally envisioned by Turing and others, did not deliver upon initial expectations. What seems possible in theory has yet to yield many viable applications in practice.

Some blame an over-commitment to symbolic systems relative to other available techniques (e.g., reinforcement learning) for the dwindling of research funding in the late 1980s known as the AI Winter. Regardless, as limitations to the capacity of “good old fashioned AI” to deliver practical applications became apparent, researchers pursued a variety of other approaches to approximating cognition grounded in the analysis and manipulation of real world data. An important consequence of the shift was that researchers began to try to solve specific problems or master particular “domains,” such as converting speech to text or playing chess, instead of pursuing a holistic intelligence capable of performing every cognitive task within one system.

All manner of AI techniques see study and use today. Much of the contemporary excitement around AI, however, flows from the enormous promise of a particular set of techniques known collectively as machine learning. Machine learning (“ML”) refers to the capacity of a system to improve its performance at a task over time. Often this task involves recognizing patterns in datasets, although ML outputs can include everything from translating languages and diagnosing precancerous moles to grasping objects or

\footnotesize
18 See STONE ET AL., supra note 7, at 51.
19 See id.
20 See id.; see also NAT’L SCI. & TECH. COUNCIL, supra note 9, at 25.
21 STONE ET AL., supra note 7, at 51.
22 Id. Originally the community drew a distinction between “weak” or “narrow” AI, designed to solve a single problem like chess, and “strong” AI with human-like capabilities across the boards. Today the term strong AI has given way to terms like artificial general intelligence (AGI), which refer to systems that can accomplish tasks in more than one domain without necessarily mastering all cognitive tasks.
23 See NAT’L SCI. & TECH. COUNCIL, supra note 9, at 8.
helping to drive a car. As alluded to above, most every technique that underpins ML has been around for decades. The recent explosion of efficacy comes from a combination of much faster computers and much more data.25

In other words, AI is an umbrella term, comprised by many different techniques. Today’s cutting-edge practitioners tend to emphasize approaches such as deep learning within ML that leverage many-layered structures to extract features from enormous data sets in service of practical tasks requiring pattern recognition, or use other techniques to similar effect.26 As we will see, these general features of contemporary AI — the shift toward practical applications, for example, and the reliance on data — also inform our policy questions.

Where is AI developed and deployed?

Development of AI is most advanced within industry, academia, and the military.27 Industry in particular is taking the lead on AI, with tech companies hiring away top scientists from universities and leveraging unparalleled access to enormous computational power and voluminous, timely data.28 This was not always the case: as with many technologies, AI had its origins in academic research catalyzed by considerable military funding.29 But industry has long held a significant role. The AI Winter gave way to the present AI Spring in part thanks to the continued efforts of researchers who once worked at Xerox Park and Bell Labs. Even today, much of the AI research occurring at firms is happening in research departments structurally insulated, to some degree, from the demands of the company’s bottom line. Still, it is worth noting that as few as seven for profit institutions — Google, Facebook, IBM, Amazon, Microsoft, Apple, and Baidu in China — seemingly hold AI capabilities that vastly outstrip all other institutions as of this writing.30

AI is deployed across a wide variety of devices and settings. How wide depends on whom you ask. Some would characterize spam filters that leverage ML or simple chat

25 STONE ET AL., supra note 7, at 51.
26 See id. at 14-15; see also NAT’L SCI. AND TECH. COUNCIL, supra note 9, at 9-10.
27 There are other, private organizations and public labs with considerable acumen in artificial intelligence, including the Allen Institute for AI and the Stanford Research Institute (SRI).
bots on social media — programmed to, for instance, reply to posts about climate change by denying its basis in science — as AI. Others would limit the term to highly complex instantiations such as the Defense Advanced Research Project Agency’s (“DARPA”) Cognitive Assistant that Learns and Organizes (“CALO”) or the guidance software of a fully driverless car. We might also draw a distinction between disembodied AI, which acquires, processes, and outputs information as data, and robotics or other cyber-physical systems, which leverage AI to act physically upon the world. Indeed, there is reason to believe the law will treat these two categories differently.

Regardless, many of the devices and services we access today — from iPhone autocorrect to Google Images — leverage trained pattern recognition systems or complex algorithms that a generous definition of AI might encompass. The discussion that follows does not assume a minimal threshold of AI complexity but focuses instead on what is different about contemporary AI from previous or constituent technologies such as computers and the Internet.

Why AI “policy”? That artificial intelligence lacks a stable, consensus definition or instantiation complicates efforts to develop an appropriate policy infrastructure. We might question the very utility of the word “policy” in describing societal efforts to channel AI in the public interest. There are other terms in circulation. A new initiative anchored by MIT’s Media Lab and Harvard University’s Berkman Klein Center for Internet and Society, for instance, refers to itself as the “Ethics and Governance of Artificial Intelligence Fund.” Perhaps these are better words. Or perhaps it makes no difference, in the end, what labels we use as long as the task is to explore and channel AI’s social impacts and our work is nuanced and rigorous.

This essay uses the term policy deliberately for several reasons. First, there are issues with the alternatives. The study and practice of ethics is of vital importance, of course, and AI presents unique and important ethical questions. Several efforts are underway, within industry, academia, and other organizations, to sort out the ethics of AI. But these efforts likely cannot substitute for policymaking. Ethics as a construct is

notoriously malleable and contested: both Kant and Bentham get to say “should.” Policy — in the sense of official policy, at least — has a degree of finality once promulgated. Moreover, even assuming moral consensus, ethics lacks a hard enforcement mechanism. A handful of companies dominate the emerging AI industry. They are going to prefer ethical standards over binding rules for the obvious reason that no tangible penalties attach to changing or disregarding ethics should the necessity arise.

Indeed, the unfolding development of a professional ethics of AI, while at one level welcome and even necessary, merits ongoing attention. History is replete with examples of new industries forming ethical codes of conduct, only to have those codes invalidated by the federal government (the Department of Justice or Federal Trade Commission) as a restraint on trade. The National Society of Professional Engineers (“NSPE”) alone has been the subject of litigation across several decades. In the 1970s, the DOJ sued the NSPE for establishing a “canon of ethics” that prohibited certain bidding practices; in the 1990s, the FTC sued the NSPE for restricting advertising practices. The ethical codes of structural engineers have also been the subject of complaints, as have the codes of numerous other industries. Will AI engineers fare differently? This is not to say companies or groups should avoid ethical principles, only that we should pay attention to the composition and motivation of the authors of such principles, as well as their likely effects on markets and on society.

The term “governance” has its attractions. Like policy, governance is a flexible term that can accommodate many modalities and structures. Perhaps too flexible: it is not entirely clear what is being governed and by whom. Regardless, governance carries its own intellectual baggage — baggage that, like “ethics,” is complicated by industry’s dominance of AI development and application. Setting aside the specific associations with “corporate governance,” much contemporary governance literature embeds the claim that authority will or should devolve to actors other than the state. While it is true

---

37 José de Sousa E Brito, Right, Duty, and Utility: From Bentham to Kant and from Mill to Aristotle, XVIII/2 REVISTA IBEROAMERICANA DE ESTUDIOS UTILITARIAS 91, 92 (2010).
39 See Hutson, supra note 34.
40 Romain Dillet, Apple Joins Amazon, Facebook, Google, IBM and Microsoft in AI Initiative, TECHCRUNCH (Jan. 27, 2017), https://techcrunch.com/2017/01/27/apple-joins-amazon-facebook-google-ibm-and-microsoft-in-ai-initiative. My own interactions with the Partnership on AI, which has a diverse board of industry and civil society, suggests that participants are genuinely interested in channeling AI toward the social good.
43 Brian Cheffins, The History of Corporate Governance, in THE OXFORD HANDBOOK CORP. GOVERNANCE 46 (Mike Wright et al. eds., 2013).
that invoking the term governance can help insulate technologies from overt government interference — as in the case of Internet governance through non-governmental bodies such as ICANN and the IETF\textsuperscript{45} — the governance model also resists official policy by tacitly devolving responsibility to industry from the state.\textsuperscript{46}

Meanwhile, several aspects of policy recommend it. Policy admits of the possibility of new laws, but does not require them. It may not be wise or even feasible to pass general laws about artificial intelligence at this early stage. Whereas it is very likely wise and timely to plan for AI’s effects on society, including through the development of expertise, the investigation of AI’s current and likely social impacts, and perhaps smaller changes to appropriate doctrines and laws in response to AI’s positive and negative affordances.\textsuperscript{47}

Industry may seek to influence public policy, but it is not its role ultimately to set it. Policy conveys the necessity of exploration and planning, the finality of law, and the primacy of public interest without definitely endorsing or rejecting regulatory intervention. For these reasons, I have consciously chosen it as my frame.

II. **KEY QUESTIONS FOR AI POLICY**

This Part turns to the main goal of the essay: a roadmap to the various challenges that AI poses for policymakers. It starts with discrete challenges, in the sense of specific domains where attention is warranted, and then discusses some general questions that tend to cut across domains. For the most part, the essay avoids getting into detail about specific laws or doctrines that require reexamination and instead emphasize questions of overall strategy and planning.

The primary purpose of this Part is to give newer entrants to the AI policy world — whether from government, industry, media, academia, or otherwise — a general sense of what kinds of questions the community is asking and why. A secondary purpose is to help bring cohesion to this multifaceted and growing field. The inventory hopes to provide a roadmap for individuals and institutions to the various policy questions that arguably require their attention. The essay tees up questions; it does not purport to answer them.

A limitation of virtually any taxonomic approach is the need to articulate criteria for inclusion — why are some questions on this list and not others?\textsuperscript{48} Experts may vary on

\textsuperscript{44} See supra note 44.

\textsuperscript{45} The United States government stood up both ICANN and IETF but today they run largely interdependent of state control as non-profits.

\textsuperscript{46} See, e.g., Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. (forthcoming 2018) (arguing inter alia for a clarification that companies may not invoke trade secret law to avoid scrutiny of their AI or algorithmic systems by criminal defendants).

\textsuperscript{47} Cf. Ryan Calo, *The Boundaries of Privacy Harm*, 86 IND. L.J. 1132, 1139–42 (2011) (critiquing Daniel Solove’s taxonomy of privacy). If I have an articulable criterion for inclusion, it is sustained attention by academics and policymakers. Some version of the questions in this Part appear in the social scientific literature, in the White House reports on AI, in the Stanford AI 100 report, in the latest U.S. Robotics
the stops they would include in a roadmap of key policy issues, and I welcome critique. There are several places where I draw distinctions or parallels that are not represented elsewhere in the literature, with which others may disagree. Ultimately this represents but one informed scholar’s take on a complex and dynamic area of study.

A. Justice and Equity

Perhaps the most visible and developed area of AI policy to date involves the capacity of algorithms or trained systems to reflect human values such as fairness, accountability, and transparency (“FAT”). This topic is the subject of considerable study, including an established but accelerating literature on technological due process and at least one annual conference on the design of FAT systems. The topic is also potentially quite broad, encompassing both the prospect of bias in AI-enabled features or products as well as the use of AI in making material decisions regarding financial, health, and even liberty outcomes. In service of teasing out specific policy issues, the essay separates “applied inequality” from “consequential decision-making” while acknowledging the considerable overlap.

Inequality in application

By inequality in application, I mean to refer to a particular set of problems involving the design and deployment of AI that works well for everyone. The examples here include everything from a camera that cautions against taking a Taiwanese-American blogger’s picture because the software believes she is blinking, to an image recognition system that characterizes an African American couple as gorillas, to a translation engine that associates the role of engineer with being male and the role of nurse with being female. These scenarios can be policy relevant in their own right, as when African Americans fail to see opportunities on Facebook due to the platform’s (now discontinued) discriminatory affordances, or Asian Americans pay more for test preparation due to a price


discriminatory algorithm. They can also hold downstream policy ramifications, as when a person of Taiwanese descent has trouble renewing a passport, or a young woman in Turkey researching international opportunities in higher education finds only references to nursing.

There are a variety of reasons why AI systems might not work well for certain populations. For example, the designs may be using models trained on data where a particular demographic is underrepresented and hence not well reflected. More white faces in the training set of an image recognition AI means the system performs best for Caucasians. There are also systems that are selectively applied to the marginalized populations. To illustrate, police use “heat maps” that purport to predict areas of future criminal activity to determine where to patrol but in fact lead to disproportionate harassment of African Americans. Yet police do not routinely turn such techniques inward to predict which officers are likely to engage in excessive force. Nor do investment firms initiate transactions on the basis of machine learning that they cannot explain to wealthy, sophisticated investors.

The policy questions here are at least twofold. First, what constitutes best practice in minimizing discriminatory bias and by what mechanism (antidiscrimination laws, consumer protection, industry standards) does society incentivize development and adoption of best practice? And second, how do we ensure that the risks and benefits of artificial intelligence are evenly distributed across society? Each set of questions is already occupying considerable resources and attention, including within the industries that build AI into their products, and yet few would dispute we have a long way to go before resolving them.

Consequential decision-making

57 See Adam Hadhazy, Biased Bots: Artificial-intelligence Systems Echo Human Prejudices, PRINCETON UNIV. (Apr. 18, 2017), https://www.princeton.edu/news/2017/04/18/biased-bots-artificial-intelligence-systems-echo-human-prejudices (last visited Sept. 14, 2017) ("Turkish uses a gender-neutral, third person pronoun, ‘o.’ Plugged into the online translation service Google Translate, however, the Turkish sentences ‘o bir doktor’ and ‘o bir hemşire’ are translated into English as ‘he is a doctor’ and ‘she is a nurse.’"). See generally Caliskan et al., supra note 53.
58 See Rose, supra note 51 (discussing performance and race in the context of camera software).
Closely related, but distinct in my view, is the question of how to design systems that make or help make consequential decisions about people. The question is distinct from unequal application in general in that consequential decision-making, especially by government, often takes place against a backdrop of procedural rules or other guarantees of process. For example, in the United States, the Constitution guarantees due process and equal protection by the government, and European Union citizens have the right to request that consequential decisions by private firms involve a human (current) as well as a right of explanation for adverse decisions by a machine (pending). Despite these representations, participants in the criminal justice system are already using algorithms to determine whom to police, whom to parole, and how long a defendant should stay in prison.

There are three distinct facets to a thorough exploration of the role of AI in consequential decision-making. The first involves cataloguing the objectives and values that procedure and process are trying to advance in a particular context. Without a thorough understanding of what it is that laws, norms, and other safeguards are trying to achieve, we cannot assess whether existing systems are adequate let alone design new systems that are. This task is further complicated by the tradeoffs and tensions inherent in such safeguards, as when the Federal Rules of Civil Procedure call simultaneously for a “just, speedy, and inexpensive” proceeding or where the Sixth Amendment lays out labor-intensive conditions for a fair criminal trial that also has to occur quickly.

The second facet involves determining which of these objectives and values can and should be imported into the context of machines. Deep learning, as a technique, may be effective in establishing correlation but unable to yield or articulate a causal

---


67 See generally Citron, *supra* note 63 (discussing the goals of technological due process); Crawford & Schultz, *supra* note 64 (discussing due process and Big Data); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2017) (arguing that current decision-making processes have not kept up with technology).

68 FED. R. CVT. P. 1. I owe this point to my colleague Elizabeth Porter.

69 U.S. CONST. amend. VI (requiring that a defendant be allowed to be presented with nature and cause of the accusations, to be confronted with the witnesses against him, to compel favorable witnesses and to have the assistance of counsel, all as part of a speedy and public trial).
AI here can say what will happen but not why. If so, the outputs of multi-layer neural nets may be inappropriate affiants for warrants, bad witnesses in court, or poor bases for judicial determinations of fact. Notions such as prosecutorial discretion, the rule of lenity, and executive pardon may not admit of mechanization at all. Certain decisions, such as the decision to take an individual off of life support, raise fundamental concerns over human dignity and thus perhaps cannot be made even by objectively well-designed machines.

A third facet involves the design and vetting of consequential decision-making systems in practice. There is widespread consensus that such systems should be fair, accountable, and transparent. However, other values — such as efficiency — are less well developed. The overall efficiency of an AI-enabled justice system, as distinct from its fairness or accuracy in the individual case, constitutes an important omission. As the saying goes, “justice delayed is justice denied”: we should not aim as a society to hold a perfectly fair, accountable, and transparent process for only a handful of people a year.

Interestingly, the value tensions inherent in processual guarantees seem to find analogs, if imperfect ones, in the machine learning literature around performance tradeoffs. Several researchers have measured how making a system more transparent or less biased can decrease its accuracy overall. More obvious than efficiency, accuracy is an important dimension of fairness: we would not think of rolling a die to determine sentence length as fair, even if it is transparent to participants and unbiased as to demographics. The policy challenge involves how to manage these tradeoffs, either by designing techno-social systems that somehow maximize for all values, or by embracing a particular tradeoff in a way society is prepared to recognize as valid. The end game of designing systems that reflect justice and equity will involve very considerable, interdisciplinary efforts and is likely to prove a defining policy issue of our time.

A. Use of Force

---


72 The rule of lenity requires courts to construe criminal statutes narrowly, even where legislative intent appeared to militate toward a broader reading. E.g., McBoyle v. United States, 283 U.S. 25, 27 (1931) (declining to extend a stolen vehicle statute to stolen airplanes). For an example of a discussion of the limits of translating laws into machine code, see Harry Surden & Mary-Anne Williams, Technological Opacity, Predictability, and Self-Driving Cars, 38 Cardozo L. Rev. 121, 162-63 (2016).

73 See James H. Moor, Are There Decisions Computers Should Never Make?, in 1 Nature and System 217, 226 (1979). This concern is also reflected in Part II.B concerning the use of force.

74 See supra notes 49–50 and accompanying text.


76 See id. at 1.
A special case of AI-enabled decision-making involves the decision to use force. As alluded to above, there are decisions — particularly involving the deliberate taking of life — that policymakers may decide never to commit exclusively to machines. Such is the gist of many debates regarding the development and deployment of autonomous weapons. International consensus holds that people should never give up “meaningful human control” over a kill decision. Yet debate lingers as to the meaning and scope of meaningful human control. Is monitoring enough? Target selection? And does the prescription extend to defensive systems as well, or only offensive tactics and weapons? None of these important questions appear settled.

There is also the question of who bears responsibility for the choices of machines. The automation of weapons may seem desirable in some circumstances or even inevitable. It seems unlikely, for example, that the United States military would permit its military rivals to have faster or more flexible response capabilities than its own whatever their control mechanism. Regardless, establishing a consensus around meaningful human control would not obviate all inquiry into responsibility in the event of mistake or war crime. Some uses of AI presuppose human decision but nevertheless implicate deep questions of policy and ethics — as when the intelligence community leverages algorithms to select targets for remotely operated drone strikes. And there are concerns that soldiers will be placed into the loop for the sole purpose of absorbing liability for wrongdoing, as anthropologist Madeline Clare Elish argues. Thus, policymakers must work toward a framework for responsibility around AI and force that is fair and satisfactory to all stakeholders.

---

77 Note that force is deployed in more contexts than military conflict. We might also ask after the propriety of the domestic use of force by border patrols, police, or even private security guards. For a discussion, see Elizabeth E. Joh, Policing Police Robots, 64 UCLA L. REV. DISCOURSE 516, 530–42 (2016).


80 Kenneth Anderson and Matthew Waxman in particular have made important contributions to the realpolitik of AI weapons. See, e.g., Kenneth Anderson & Matthew Waxman, Law and Ethics for Autonomous Weapon Systems: Why a Ban Won’t Work and How the Laws of War Can, TASK FORCE ON NAT’L SECURITY & L. (Hoover Institution, Stanford U.), (2013), at 2 (arguing that automated weapons are both desirable and inevitable).

81 See id.

82 See John Naughton, Death by Drone Strike, Dished Out by Algorithm, GUARDIAN (Feb. 21, 2016, 3:59 AM EST), https://www.theguardian.com/commentisfree/2016/feb/21/death-from-above-nia-csa-skynet-algorithm-drones-pakistan (“General Michael Hayden, a former director of both the CIA and the NSA, said this: ‘We kill people based on metadata’.”).

B. Safety and Certification

As the preceding section demonstrates, AI systems do more than process information and assist officials to make decisions of consequence. Many systems — such as the software that controls an airplane on autopilot or a fully driverless car — exert direct and physical control over objects in the human environment. Others provide sensitive services that, when performed by people, require training and certification. These applications raise additional questions concerning the standards to which AI systems are held and the procedures and techniques available to ensure those standards are being met.84

Setting and validating safety thresholds

Robots and other cyber-physical systems have to be safe. The question is how safe, and how do we know. In a wide variety of contexts, from commercial aviation to food safety, regulatory agencies set specific safety standards and lay out requirements for how those standards must be met. Such requirements do not exist for many robots.

Members of Congress and others have argued that we should embrace, for instance, driverless cars, to the extent that robots are or become safer drivers than humans.85 But “safer than humans” seems like an inadequate standard by which to vet any given autonomous system. Must the system be safer than humans unaided or humans assisted by cutting-edge safety features? Must the system be safer than humans overall or across all driving conditions? And just how much safer must driverless cars be than people before we tolerate or incentivize them? These are ultimately difficult questions not of technology but of policy.86

Even assuming policymakers set satisfactory safety thresholds for driverless cars, drone delivery, and other instantiations of AI, we need to determine a proper and acceptable means of verifying that these standards are met. This process has an institutional or “who” component — e.g., government testing, third-party independent certification, and self-certification by industry — as well as a technical or “how” component — e.g., unit testing, fault-injection, virtualization, supervision.87 Local and international standards can be a starting point, but considerable work remains — especially as new potential applications and settings arise. For example, we might resolve safety thresholds for drone delivery or warehouse retrieval only to revisit the question anew for sidewalk delivery and fast food preparation.

84 See HENRIK I. CHRISTENSEN ET AL., FROM INTERNET TO ROBOTICS 105-09 (Nov. 7, 2016), http://jacobsschool.ucsd.edu/contextualrobotics/docs/rm3-final-rs.pdf; Stone et al., supra note 7, at 42.
87 Cf. Bryant Walker Smith, How Governments Can Promote Automated Driving, 47 N.M. L. REV. 99, 101 (2017) (discussing different avenues through which government can promote automated driving and prepare community conditions to facilitate seamless integration of driverless cars once they become road-worthy).
There are further complications still. Some systems, such as high speed trading algorithms that can destabilize the stock market or cognitive radio systems that can interfere with emergency communications, may hold the potential, alone or in combination, to cause serious indirect harm. Others may engage in harmful acts such as disinformation that simultaneously implicate free speech concerns. Policymakers must determine what kinds of non-physical or indirect harms rise to the level that regulatory standards are required. Courts have a role in setting safety policy in the United States though the imposition of liability. It turns out that AI — especially AI that displays emergent properties — may pose challenges for civil liability. Courts or regulators must address this misalignment. Markets have a role as well, for instance, through the availability and conditions of insurance.

**Certification**

A closely related policy question arises where AI performs a task that, when done by a human, requires evidence of specialized skill or training. In some contexts, society has seemed comfortable thus far dispensing with the formal requirement of certification when technology can be shown to be capable through supervised use. This is true of the autopilot modes of airplanes, which do not have to attend flight school. The question is open with respect to vehicles.

But what of technology under development today, such as autonomous surgical robots, the very value of which turns on bringing skills into an environment where no one has them? And how do we think about systems that purport to dispense legal, health, or financial advice, which requires adherence to complex fiduciary and other duties pegged to human judgment? Surgeons and lawyers must complete medical or law school and pass boards or bars. This approach may or may not serve an environment rich in AI, a dynamic that is already unfolding as the Food and Drug Administration works to classify downloadable mobile apps as medical devices and other apps to dispute parking tickets.

---

90 See Calo, supra note 33, at 538-45.
91 For an overview, see Andrea Bertolini et al., On Robots and Insurance, 8 INT’L J. SOC. ROBOTICS 381, 381 (2016) (discussing the need for adaptations in the insurance industry to respond to robotics).
92 Christensen et al., supra note 84, at 105.
93 See Mark Harris, Will You Need a New License to Operate a Self-Driving Car?, IEEE SPECTRUM (Mar. 2, 2015, 3:00 PM GMT), https://spectrum.ieee.org/cars-that-think/transportation/human-factors/will-you-need-a-new-license-to-operate-a-self-driving-car (discussing the current unsettled state of licensing schemes for “passengers” of driverless cars).
Cybersecurity

Finally, it is becoming increasingly clear that AI complicates an already intractable cybersecurity landscape. First, as alluded to above, AI increasingly acts directly and even physically on the world. When a malicious party gains access to a cyber-physical system, suddenly bones instead of bits are on the line. Second, ML and other AI techniques have the potential to alter both the offensive and defensive capabilities around cybersecurity, as dramatized by a recent competition held by DARPA where AI agents attacked and defended a network autonomously. AI itself creates a new attack surface in the sense that ML and other techniques can be coopted purposefully to trick the system — an area known as adversarial machine learning. New threat models, standards, and techniques must be developed to address the new challenges of securing information and physical infrastructures.

C. Privacy and Power

Over the past decade, the discourse around privacy has shifted perceptibly. What started out as a conversation about individual control over personal information has evolved into a conversation around the power of information more generally, i.e., the control institutions have over consumers and citizens by virtue of possessing so much information about them. The acceleration of artificial intelligence, which is intimately tied to the availability of data, will play a significant role in this evolving conversation in at least two ways: (1) the problem of pattern recognition and (2) the problem of data parity. Note that unlike some of the policy questions discussed above, which envision the consequential deployment of imperfect AI, the privacy questions that follow assume AI that is performing its assigned tasks only too well.

The Problem of Pattern Recognition

———

96 See generally GREG ALLEN & TANIEL CHAN, BELFER CTR. FOR SCI. & INT’L AFFAIRS, ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY (2017) (discussing ways of advancing policy on AI and national security).
97 See supra Part II.B.
98 See M. Ryan Calo, Open Robotics, 70 Md. L. Rev. 571, 593-601 (2011) (discussing how robots have the ability to cause physical damage and injury).
100 The flagship privacy law workshop — Privacy Law Scholars Conference — recently celebrated its tenth anniversary, although of course privacy discourse goes back much further.
The capacity of AI to recognize patterns people cannot themselves detect threatens to eviscerate the already unstable boundary between what is public and what is private. Artificial intelligence is increasingly able to derive the intimate from the available. This means that freely shared information of seeming innocence — where you ate lunch, for example, or what you bought at the grocery store — can lead to insights of a deeply sensitive nature. With enough data about you and the population at large, firms, governments, and other institutions with access to AI will one day make guesses about you that you cannot imagine — what you like, whom you love, what you have done.

Several serious policy challenges follow. The first set of challenges involves the acceleration of an existing trend around information extraction. Consumers will have next to no ability to appreciate the consequences of sharing information. This is a well-understood problem in privacy scholarship. The community has addressed these challenges to privacy management under several labels, from databases to big data. In that the entire purpose of AI is to spot patterns people cannot, however, the issue is rapidly coming to a head. Perhaps the mainstreaming of AI technology will increase the pressure on policymakers to step in and protect consumers. Perhaps not. Researchers are, at any rate, already exploring various alternatives to the status quo: fighting fire with fire by putting AI in the hands of consumers, for example, or abandoning notice and choice altogether in favor of rules and standards. Whatever path we take should bear in mind the many ways powerful firms can subvert and end run consumer interventions and the unlikelihood consumers will keep up in a technological arms race.

Consumer privacy is under siege. Citizens, meanwhile, will have next to no ability to resist or reform surveillance. Two doctrines in particular interact poorly with the new affordances of artificial intelligence, both related to the reasonable expectation of privacy standard embedded in American constitutional law. First, the interpretation of the Fourth Amendment by the courts that citizens enjoy no reasonable expectation of privacy in

---

106 For example, Decide.com was an artificially intelligent tool to help consumer decide when to purchase products and services. Decide.com was eventually acquired by eBay. John Cook, eBay Acquires Decide.com, GEEKWIRE (Sep. 6, 2013), https://www.geekwire.com/2013/ebay-acquires-decidecom-shopping-research-site-shut-sept-30/.
public or from a public vantage does not seem long for this world. If everyone in public can be identified through facial recognition, and if the “public” habits of individuals or groups permit AI to derive private facts, then citizens will have little choice but to convey information to a government bent on public surveillance. Second, and related, the interpretation by the courts that individuals have no reasonable expectation of privacy in (non-content) information they convey to a third party such as the telephone company will continue to come under strain.

Here is an area where grappling with legal doctrine seems inevitable. Courts are policymakers of a kind and the judiciary is already responding to these new realities by requiring warrants or probable cause in contexts involving public movements or third party information. For example, in United States v. Jones, the Supreme Court required a warrant for officers to affix a GPS to a defendant’s vehicle for the purpose of continuous monitoring. Five Justices in Jones articulated a concern over law enforcement’s ability to derive intimate information from public travel over time. There is a case before the Court as of this writing concerning the ability of police to demand historic location data about citizens from their mobile phone provider.

On the other hand, in the dog-sniffing case Florida v. Jardines, the Court also reaffirmed the principle that individuals have no reasonable expectation of privacy in contraband such as illegal drugs. Thus, in theory, even if the courts resolve to recognize a reasonable expectation of privacy in public and in information conveyed to a third party, courts might still permit the government to leverage AI to search exclusively for illegal activity. Indeed, some argue that AI is not a search at all given that no human need to access the data unless or until the AI identifies something unlawful. Even assuming away the likely false positives, a reasonable question for law and policy is whether we want to live in a society with perfect enforcement.

The second set of policy challenges involves not what information states and firms collect but the way highly granular information gets deployed. Again, the privacy conversation has evolved to focus not on the capacity of the individual to protect her data, but on the power over an individual or group that comes from knowing so much about them. For example, firms can manipulate other market participants through a fine-tuned understanding of the individual and collective cognitive limitations of consumers. Bots

---

109 Courts and statutes tend to recognize that the content of a message such as an email deserves greater protection than the non-content that accompanies the message, e.g., where it is going, whether it is encrypted, whether it contains attachments, and so on. Cf. Riley v. California, 134 S. Ct. 2473 (2014) (invalidating the warrantless search and seizure of a mobile phone incident to arrest).
111 Carpenter v. United States of America, 819 F.3d 880, 886 (6th Cir. 2016), cert. granted, 137 S. Ct. 2211 (June 5, 2017) (No. 16-402).
can gain our confidences to extract personal information. Politicians and political operatives can micro-target messages, including misleading ones, in an effort to sway aggregate public attention. All of these capacities are dramatically enhanced by the ability of AI to detect patterns in a complex world. Thus, a distinct area of study is the best law and policy infrastructure for a world of such exquisite and hyper-targeted control.

The Data Parity Problem

The data-intensive nature of machine learning, the technique yielding the most powerful applications of AI at the moment, has ramifications that are distinct from the pattern recognition problem. Simply put, the greater access to data a firm has, the better positioned it is to solve difficult problems with ML. As Amanda Levendowski explores, ML practitioners have essentially three options in securing sufficient data. They can build the databases themselves, they can buy the data, or they can use “low friction” alternatives such as content in the public domain. The last option carries perils for bias discussed above. The first two are avenues largely available to big firms or institutions such as Facebook or the military.

The reality that a handful of large entities (literally, fewer than a human has fingers) possess orders of magnitude more data than anyone else leads to a policy question around data parity. Smaller firms will have trouble entering and competing in the marketplace. Industry research labs will come to outstrip public labs or universities, to the extent they do not already. Accordingly, cutting-edge AI practitioners will face even greater incentives to enter the private sphere, and ML applications will bend systematically toward the goals of profit-driven companies and not society at large. Companies will possess not only more and better information but a monopoly on its serious analysis.

Why label the question of asymmetric access to data a “privacy” question? I do so because privacy ultimately governs the set of responsible policy outcomes that arise in response to the data parity problem. Firms will, and already do, invoke consumer privacy as a rationale for not permitting access to their data. This is partly why the AI policy community must maintain a healthy dose of skepticism toward “ethical codes of conduct” developed by industry. Such codes are likely to contain a principle of privacy that, unless carefully crafted, operates to help shield the company from an obligation to share training data with other stakeholders.

119 See id.
120 See id.
121 See supra PART I.
A related question involves access to citizen data held by the government. Governments possess an immense amount of information; data that citizens are obligated to provide to the state forms the backbone of the contemporary data broker industry. Firms big and small, as well as university and other researchers, may be able to access government data on comparable terms. But there are policy challenges here as well. Governments can and should place limits and conditions around sharing data.

In the United States at least, this means carefully crafting policies to avoid constitutional scrutiny as infringements on speech. The government cannot pick and choose with impunity the sorts of uses to which private actors place data released by the state. At the same time, governments may be able to put sensible restrictions in place before compelling citizens to release private data.

To be clear: I do not think society should run roughshod over privacy in its pursuit of data parity. Indeed, I present this issue as a key policy challenge precisely because I believe we need mechanisms by which to achieve a greater measure of data parity without sacrificing personal or collective privacy. Some within academia and industry are already working on methods — including differential privacy and federated training — that seek to minimize the privacy impact of granting broader access to data-intensive systems. The hard policy question is how to incentivize technical, legal, social, and other interventions that safeguard privacy even as AI is democratized.

A. Taxation and Displacement of Labor

A common concern, especially in public discourse, is that AI will displace jobs by mastering tasks currently performed by people. The classic example is the truck driver: many have observed that self-driving vehicles could obviate, or at least radically transform, this very common role. Machines have been replacing people since the Industrial Revolution (which was hard enough on society). The difference, many suppose, is twofold: first, the process of automation will be much faster, and second, very few sectors will remain untouched by AI’s contemporary and anticipated capabilities. This would widen the populations that could feel AI’s impact and limits the efficacy of unemployment benefits or retraining.

126 See, e.g., Ford, supra note Error! Bookmark not defined. (“[M]achines themselves are turning into workers.”).
In its exploration of AI’s impact on America, the Obama White House specifically inquired into the impact of AI on the job force and issued a report recommending a thicker social safety net to manage the upcoming disruption. Some predict that new jobs will arise even as old ones fall away, or that AI will often improve the day to day of workers by permitting them to focus on more rewarding tasks involving judgment and creativity with which AI struggles. Others explore the eventual need for a universal basic income, presumably underwritten by gains in productivity for automation, so that even those displaced entirely by AI have access to resources. Still others wisely call for more and better information specific to automation so as to be able better to predict and scope the effects of AI.

In addition to assessing impact and addressing displacement, policymakers will have to think through the effects of AI on the public fisc. Taxation is a highly complex policy domain that touches upon virtually all aspects of society; AI is no exception. Robots do not pay taxes, as the IRS once remarked in letter. Bill Gates, Jr. thinks they should. Others warn that a tax on automation amounts to a tax on innovation and progress. Ultimately, federal and state policymakers will have to figure out how to keep the lights on in the absence of, for instance, the bulk of today’s income taxes.

B. Cross-Cutting Questions (Selected)

The preceding list of questions is scarcely exhaustive as to consequences of artificial intelligence for law and policy. Notably missing is any systemic review of the ways AI challenges existing legal doctrines. For example, that AIs are capable of generating spontaneous speech or content raises doctrinal questions around the limits of the First Amendment as well as the contours of intellectual property. Below, this essay discusses the prospect that AI will wake up and kill us, which, if true, would seem to

render every other policy context moot. But the preceding inventory does cover most of the common big picture policy questions that tend to dominant serious discourse around artificial intelligence.

In addition to these specific policy contexts such as privacy, labor, or the use of force, recurrent issues arise that cut across domains. I have selected a few here that deserve greater attention: determining the best institutional configuration for governing AI, investing collective resources in AI that benefit individuals and society, addressing hurdles to AI accountability, and addressing our tendency to anthropomorphize technologies such as AI. I will discuss each of these systemic questions briefly in turn.

**Institutional Configuration and Expertise**

The prospect that AI presents individual or systemic risk, while simultaneously promising enormous potential benefits to people and society if responsibly deployed, presents policymakers with an acute challenge around the best institutional configuration for channeling AI. Today AI policy is done, if at all, by piecemeal; federal agencies, states, cities, and other government units tackle issues that most relate to them in isolation. There are advantages to this approach similar to the advantages inherent in federalism — the approach is sensitive to differences across contexts and preserves room for experimentation. But some see the piecemeal approach as problematic, calling, for instance, for a kind of FDA for algorithms to vet every system with a serious potential to cause harm.

AI prefigures into a common, but I think misguided, observation about the relationship between law and technology. The public sees law as too slow to catch up to technologic innovation. Sometimes it is true that particular laws or regulations become long outdated as technology moves beyond where it was when the law was passed. For example, the Electronic Communications Privacy Act, passed in 1986, interacts poorly with a post Internet environment in part because of ECPA’s assumptions about how electronic communications would work. But this is hardly inevitable, and often political. The Federal Trade Commission has continued in its mission of protecting markets and consumers unabated, in part because it enforces a standard — that of unfair and deceptive practice — that is largely neutral as to technology. In other contexts, agencies have passed new rules or interpreted rules differently to address new techniques and practices.

The better-grounded observation is that government lacks the requisite expertise to manage society in such a deeply technically-mediated world. Government bodies are

---

136 See infra Part III.
137 New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) (classic articulation of concept that states serve as laboratories of democracy).
141 Calo, supra note 87.
slow to hire up and face steep competition from industry. When the state does not have
its own experts, it must either rely on the self-interested word of private firms or their
proxies or experience a paralysis of decision and action that ill-serves innovation.\textsuperscript{142} Thus, one overarching policy challenge is how best to introduce expertise about AI and robotics into all branches and levels of government so they can make better decisions with greater confidence.

The solution could involve new advisory bodies, such as an official Federal Advisory Committee on Artificial Intelligence with an existing department or a standalone Federal Robotics Commission.\textsuperscript{143} Or it could involve resuscitating the Office of Technology Assessment, building out the Congressional Research Service, or growing the Office of Science and Technology Policy. Yet another approach involves each branch hiring its own technical staff at every level. The technical knowledge and affordances of the government — from the ability to test claims in a laboratory to a working understanding of AI in lawmakers and the judiciary — will ultimately affect the government’s capacity to generate wise AI policy.

\textit{Investment and Procurement}

The government possesses a wide variety of means by which to channel AI in the public good. As recognized by the Obama White House, which published a separate report on the topic, one way to shape AI is by investing in it.\textsuperscript{144} Investment opportunities include not only basic AI research, which advance the state of computer science and help ensure the United States remains globally competitive, but also support of social scientific research into AI’s impacts on society. Policymakers can be strategic about where funds are committed and emphasize, for example, projects with an interdisciplinary research agenda and a vision for the public good.

In addition, and sometimes less well-recognized, the government can influence policy through what it decides to purchase.\textsuperscript{145} States are capable of exerting considerable market pressures. Thus, policymakers at all levels ought to be thinking about the qualities and characteristics of the AI-enabled products government will purchase and the companies that create them. Policymakers can also use contract to help ensure best practice around privacy, security, and other values. This can in turn move the entire market toward more responsible practice and benefit society overall.

\textit{Removing Hurdles to Accountability}

\textsuperscript{142} \textit{Id.} (giving examples).
\textsuperscript{145} \textit{See, e.g.}, Smith, \textit{supra} note 87 (discussing procurement in connection with driverless cars); Whittington et al., \textit{supra} note 122 (discussing procurement in connection with open municipal data).
Many AI systems in use or development today are proprietary and owners of AI systems have inadequate incentives to open them up to scrutiny. In many contexts, outside analysis is necessary for accountability. For example, in the context of justice and equity, defendants may seek to challenge adverse risk scores. Or, in the context of safety and certification, third parties seek to verify claims of safety or to evidence a lack of compliance. Several reports, briefs, and research papers have called upon policymakers to remove actual or perceived barriers to accountability, including: (1) trade secret law, (2) the Computer Fraud and Abuse Act, and (3) the anti-circumvention provision of the Digital Millennium Copyright Act. This has led a number of experts to recommend the formal policy step of planning to remove such barriers in order to foster greater accountability for AI.

Mental Models of AI

The next and final Part is devoted to a discussion of whether AI is likely to end humanity, itself partly a reflection of the special set of fears that tend to accompany anthropomorphic technology such as AI. Policymakers arguably owe it to their constituents to hold a clear and accurate mental model of AI themselves and may have a role in educating citizens about the technology and its potential effects. Here they face an uphill battle, at least in the United States, due to decades of books, films, television shows, and even plays that depict AI as a threatening substitute for people. That the task is difficult, however, does not discharge policymakers from their responsibilities.

At a more granular level, the fact that instantiations of AI such as Alexa (Echo), Siri, and Cortana, not to mention countless chat bots on a variety of social media platforms, take the form of social agents presents special challenges for policy driven by our hardwired responses to social technology as though it were human. These include the potential to influence children and other vulnerable groups in commercial settings and the prospect of disrupting civic or political discourse or the further diminution of possibilities for...
solitude through a constant sense of being in the presence of another. Others are concerned about the prospect of intimacy, including sexual, between people and machines. Whatever the particulars, that even the simplest AI can trigger social and emotional responses in people requires much more study and thought.

I. ON THE AI APOCALYPSE

Some set of readers may feel I have left out a key question: does artificial intelligence present an existential threat to humanity? If so, perhaps all other discussions constitute the policy equivalent of rearranging deck chairs on the Titanic. Why fix the human world if AI is going to end it?

My own view is that AI does not present an existential threat to humanity, at least not in anything like the foreseeable future. Further, devoting disproportionate attention and resources to the AI apocalypse has the potential to distract policymakers from addressing AI’s more immediate harms and challenges and could discourage investment in research on AI’s present social impacts. How much attention to pay to a remote but dire threat is itself a difficult question of policy. If there is any risk to humanity then if follows some thought and debate is worthwhile. But too much attention has real-world consequences.

Entrepreneur Elon Musk, physicist Stephen Hawking, and other famous individuals apparently believe AI represents civilization’s greatest threat to date. The most common citation for this proposition is the work of a British speculative philosopher named Nick Bostrom. In Superintelligence, Bostrom purports to demonstrate that we are on a path toward developing AI that is both enormously superior to human intelligence and presents a significant danger of turning on its creators. Bostrom, it should be said, does not see a malignant superintelligence as inevitable. But he presents the danger as acute enough to merit serious consideration.

A number of prominent voices in artificial intelligence have convincingly challenged Superintelligence’s thesis along several lines. First, they argue that there is simply no

156 See generally Crawford & Calo, supra note 60 (“Fears about the future impacts of artificial intelligence are distracting researchers from the real risks of deployed systems . . . .”).
158 See generally Nick Bostrom, SUPERINTELLIGENCE: PATHS, DANGERS, STRATEGIES (2014) (exploring the “most daunting challenge humanity has ever faced” assessing how we might best respond).
159 See Raffi Khatchadourian, The Doomsday Invention, THE NEW YORKER (Nov. 23, 2015), https://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom. In other work, Bostrom argues that we are likely all living in a computer simulation created by our distant descendants. Nick Bostrom, Are You Living in A Simulation? 53 PHIL. Q. 211, 211 (2003). This prior claim raises an interesting paradox: if AI kills everyone in the future, then we cannot be living in a computer simulation created by our decedents. And if we are living in a computer simulation created by our
path toward machine intelligence that rivals our own across all contexts or domains. Yes, a machine specifically designed to do so can beat any human at chess. But nothing in the current literature around ML, search, reinforcement learning, or any other aspect of AI points the way toward modeling even the intelligence of a lower mammal in full, let alone human intelligence. Some say this explains why claims of a pending AI apocalypse come almost exclusively from the ranks of individuals such as Musk, Hawking, and Bostrom who lack work experience in the field. Second, critics of the AI apocalypse argue that even if we were able eventually to create a superintelligence, there is no reason to believe it would be bent on world domination, unless this were for some reason programmed into the system. As Yann LeCun, deep learning pioneer and head of AI at Facebook colorfully puts it: computers do not have testosterone.

Note that the threat to humanity could come in several forms. The first is that AI wakes up and purposefully kills everyone out of animus or to make more room for itself. This is the stuff of Hollywood movies and books by Daniel Wilson and finds next to no support in the computer science literature (which is why we call it science fiction). The second is that AI accidentally kills everyone in the blind pursuit of some arbitrary goal — for example, an irresistibly powerful AI charged with making paperclips destroys the Earth in the process of mining for materials. Fantasy is replete with examples of this scenario as well, from The Sorcerer’s Apprentice in Disney’s Fantasia to the ill-fated King Midas who demands the wrong blessing. A third is that a very bad individual or group uses AI as part of an attempt to end human life.

Even if you believe the mainstream AI community that we are hundreds of years away from understanding how to create machines capable of formulating an intent to harm, and would not do so anyway, you might be worried about the second and third scenarios. The argument has its attractions: people can set goals for AI that lead to unintended consequences. Computers do what you tell them to do, as the saying goes, not what you decreedents, then AI didn’t kill everyone. I think it a fair deduction that Professor Bostrom is wrong about something.

---

160 See Erik Sofge, Why Artificial Intelligence Will Not Obliterate Humanity, POPULAR SCI. (Mar. 19, 2015), http://www.popsci.com/why-artificial-intelligence-will-not-obliterate-humanity. Australian computer scientist Mary Anne Williams once remarked to me, “We have been doing artificial intelligence since that term was coined in the 1950s, and today robots are about as smart as insects.”

161 See Connie Loizos, This Famous Roboticist Doesn’t Think Elon Musk Understands AI, TECHCRUNCH (July 19, 2017), https://techcrunch.com/2017/07/19/this-famous-roboticist-doesnt-think-elon-musk-understands-ai/ (quoting Rodney Brooks as noting that AI alarmists “share a common thread, in that: they don’t work in AI themselves”).

162 Dave Blanchard, Musk’s Warning Sparks Call for Regulating Artificial Intelligence, NPR (July 19, 2017), http://www.npr.org/sections/alltechconsidered/2017/07/19/537961841/musks-warning-sparks-call-for-regulating-artificial-intelligence (citing an observation by Yan LeCunn that the desire to dominate is not necessarily correlated with intelligence).


164 E.g., BOSTROM, supra note 156, at 123.

165 FANTASIA (Walt Disney Productions 1940); ARISTOTLE, POLITICS I.9.11, 1257b (B. Jowett trans., 1885) I owe the analogy to King Midas to Stuart Russell, a prominent computer scientist at UC Berkeley who is among the handful of AI experts to join Musk and others in worrying aloud about AI’s capacity to threaten humanity.
want them to do. But about it is also important to consider the characteristics of the system AI doomsayers envision. It is simultaneously so primitive as to perceive a singular goal, such as making paperclips, arbitrarily assigned by a person, and yet so advanced as to be capable of outwitting and overpowering the sum total of humanity in pursuit of this goal. I find this combination of qualities unlikely, perhaps on par with the likelihood of a malicious AI bent on world domination.

Perhaps more worrying is the potential that a person or group might use AI in some way to threaten all of society. This is the vision of, for example, Daniel Suarez in his book *Daemon* and has been explored by workshops such as *Bad Actors in AI* at Oxford University. We can imagine, for example, a malicious actor leveraging AI to compromise nuclear security, using trading algorithms to destabilize the market, or spreading misinformation through AI-enabled micro-targeting to incite violence. The path from malicious activity to existential threat, however, is narrow, and for now the stuff of graphic novels.

Only time can tell us for certain who is wrong and who is right. Although it may not be the mainstream view among AI researcher and practitioners, I have attended several events where established computer scientists and other smart people reflected some version of the doomsday scenario. If there is even a remote chance that AI will wake up and kill us, i.e., if the AI apocalypse is a low probability, high loss problem, then perhaps we should pay some attention to the issue.

The strongest argument against focusing overly on Skynet or HAL in 2017 is the opportunity cost. AI presents numerous pressing challenges to individuals and society in the very short term. The problem is not that artificial intelligence “will get too smart and take over the world,” computer scientist Pedro Domingos writes, “the real problem is that [it’s] too stupid and [has] already.” By focusing so much energy on a quixotic existential threat, we risk, in information scientist Solon Barocas’ words, an AI Policy Winter.

**Conclusion**

This essay had two goals. First, it sought to provide a brief primer on artificial intelligence by defining AI in relation to previous and constituent technologies and by noting the ways the contemporary conversation around AI may be unique. One of the most obvious breaks with the past is the extent and sophistication of the policy response to AI in the United States and around the world. Thus the essay sought, second, to

---

provide an inventory or roadmap of the serious policy questions that have arisen to date. The purpose of this inventory is to inform AI policymaking, broadly understood, by identifying the issues and developing the questions to the point that readers can initiate their own investigation. The roadmap is idiosyncratic to the author but informed by longstanding participation in AI policy.

AI is remaking aspects of society today and likely to shepherd in much greater changes in the coming years. As this essay emphasized, the process of societal transformation carries with it many distinct and difficult questions of policy. Even so, there is reason for hope. We have certain advantages over our predecessors. The previous industrial revolutions had their lessons and we have access today to many more policymaking bodies and tools. We have also made interdisciplinary collaboration much more of a standard practice. But perhaps the greatest advantage is timing: AI has managed to capture policymakers’ imaginations early enough in its life-cycle that there is hope we can yet channel it toward the public interest. I hope this essay contributes in some small way to this process.
MACHINES WITHOUT PRINCIPALS: LIABILITY RULES AND ARTIFICIAL INTELLIGENCE

INTRODUCTION

The idea that humans could, at some point, develop machines that actually “think” for themselves and act autonomously has been embedded in our literature and culture since the beginning of civilization. But these ideas were generally thought to be religious expressions--what one scholar describes as an effort to forge our own Gods--or pure science fiction. There was one important thread that tied together these visions of a special breed of superhuman men/machines: They invariably were stronger, smarter, and sharper analytically; that is, superior in all respects to humans, except for those traits involving emotional intelligence and empathy. But science fiction writers were of two minds about the capacity of super-smart machines to make life better for humans.

One vision was uncritically Utopian. Intelligent machines, this account goes, would transform and enlighten society by performing the mundane, mind-numbing work that keeps humans from pursuing higher intellectual, spiritual, and artistic callings. This view was captured in the popular animated 1960s television show The Jetsons. As its title suggests, the show's vision is decidedly futuristic. The main character, George Jetson, lives with his family in a roomy, bright, and lavishly furnished apartment that seems to float in the sky. George and his family travel in a flying saucer-like car that drives itself and folds into a small briefcase. All of the family's domestic needs are taken care of by Rosie, the robotic family maid and housekeeper, who does the household chores and much of the parenting. George does “work.” He is employed as a “digital index operator” by Spacely's Space Sprockets, which makes high tech equipment. George often complains of overwork, even though he appears to simply push buttons on a computer for three hours a day, three days a week. In other words, the Jetsons live the American dream of the future.

In tangible ways, this Utopian vision of the partnership between humans and highly intelligent machines is being realized. Today, supercomputers can beat humans at their own games. IBM's “Deep Blue” can beat the pants off chess grand-masters, while its sister-super-computer “Watson” can clobber the reigning Jeopardy champions. But intelligent machines are more than show. Highly sophisticated robots and other intelligent machines perform critical functions that not long ago were thought to be within the exclusive province of humans. They pilot sophisticated aircraft; perform delicate surgery; study the landscape of Mars; and through smart nanotechnology, microscopic machines may soon
deliver targeted medicines to areas within the body that are otherwise unreachable. In every one of these examples, machines perform these complex and at times dangerous tasks as well as, if not better than, humans.

But science fiction writers also laid out a darker vision of intelligent machines and feared that, at some point, autonomously thinking machines would turn on humans. Some of the best science fiction expresses this dystopian view, including Stanley Kubrick's 1968 classic film 2001: A Space Odyssey. The film's star is not the main character, “Dave” (Dr. David Bowman, played by Keir Dullea), or “Frank” (Dr. Frank Poole, played by Gary Lockwood), who are astronauts on a secret and mysterious mission to Jupiter. Instead, the character who rivets our attention is HAL 9000, the all-knowing supercomputer who controls most of the ship's operations, but does so under the nominal command of the astronauts. The complexity of the relationship between man and the super-intelligent machine is revealed early in the film. During a pre-mission interview, HAL claims that he is “foolproof and incapable of error,” displaying human-like hubris. And when Dave is asked if HAL has genuine emotions, he replies that HAL appears to, but that the truth is unknown.

Once the mission begins, tensions between HAL and the astronauts start to surface. HAL wants the astronauts to tell him the details of the highly secret mission, but Dave and Frank refuse. In fact, they too do not know. Soon thereafter, HAL warns of the impending failure of a critical antenna on the spaceship's exterior. Starting to have doubts about HAL, Dave and Frank lock themselves in an evacuation vehicle to ensure that HAL cannot overhear their conversation; HAL reads their lips through the vehicle's window. Dave vainly tries to rescue Frank, and as soon as Dave leaves the spacecraft, HAL turns off the life-support system for the three remaining crew members, who were in suspended animation. HAL then refuses to let Dave back onto the spaceship, telling Dave that the plan to deactivate him jeopardizes the mission. Ultimately, Dave makes his way back onto the spaceship and starts shutting HAL down. All the while, as HAL regresses, HAL pleads with Dave to stop, and finally expresses fear of his demise.

The question one might ask at this point is what relevance does 2001: A Space Odyssey have to liability rules for autonomous thinking machines? The answer is quite a bit. Today's machines, as path-breaking as they are, all have a common feature that is critical in assessing liability. In each case, the machine functions and makes decisions in ways that can be traced directly back to the design, programming, and knowledge humans embedded in the machine. The human hand defines, guides, and ultimately controls the process, either directly or because of the capacity to override the machine and seize control. As sophisticated as these machines are, they are, at most, semi-autonomous. They are tools, albeit remarkably sophisticated tools, used by humans.

Where the hand of human involvement in machine decision-making is so evident, there is no need to reexamine liability rules. Any human (or corporate entity that has the power to do things that humans do, enter into contracts, hire workers, and so forth) that has a role in the development of the machine and helps map out its decision-making is potentially responsible for wrongful acts—negligent or intentional—committed by, or involving, the machine. The reason, of course, is that these machines, notwithstanding their sophistication, have no attribute of legal personhood. They are agents or instruments of other entities that have legal capacity as individuals, corporations, or other legal “persons” that may be held accountable under the law for their actions.

But the fully autonomous machines that at some point will be introduced into the marketplace may be quite different, and for that reason, society will need to consider whether existing liability rules will be up to the task of assigning responsibility.
for any wrongful acts they commit. The first generation of fully autonomous machines--perhaps driver-less cars and fully independent drone aircraft--will have the capacity to act completely autonomously. They will not be tools used by humans; they will be machines deployed by humans that will act independently of direct human instruction, based on information the machine itself acquires and analyzes, and will often make highly consequential decisions in circumstances that may not be anticipated by, let alone directly addressed by, the machine's creators. Artificial intelligence theorists distill the concept of full autonomy down to the paradigm of machines that “sense-think-act” without human involvement or intervention. And Oxford Professor Nick Bostrom, an eminent futurist, goes as far as to suggest that machines “capable of independent initiative and of making their own plans . . . are perhaps more appropriately viewed as persons than machines.”

Assuming that this description of the capabilities of such machines is accurate, the key conceptual question that autonomous thinking machines will pose is whether it is fair to think of them as agents of some other individual or entity, or whether the legal system will need to decide liability issues on a basis other than agency. To be sure, it is hard to conceptualize a machine as being anything other than an agent of a person, be it a real person or an entity with legal personhood. But there is another argument that is worth exploring, namely that concepts of agency may be frayed, if not obliterated, by autonomous thinking machines, even those that are not truly “sentient.” Let us go back to HAL. At some point before he turns murderous, HAL became an “agent” of no one. An agent who decides to go on his own frolic and detour, defying the instructions of his principal, is no longer an agent under any conventional understanding of the law. And HAL plainly detoured. HAL was given the ability to think and act independently, so much so that he “decided” to violate the first rule of robotics: That is, machines must do no harm to humans or to humanity. By deciding to harm humans, HAL at least arguably (if not decisively) terminated his status as an agent.

To be sure, HAL's capacity to violate that rule could be attributable to a manufacturing defect, design flaw, or the result of poor programming. In that case, existing legal principles would almost certainly place liability on the shoulders of HAL's creators. But suppose HAL's wrongful conduct was not a manufacturing, design, or programming flaw, but was instead an unforeseen byproduct of teaching machines to “think.” Sentient beings often choose to break rules, and all disciplines interested in artificial intelligence--technologists, scientists, ethicists, and philosophers alike--have long worried that giving machines the capacity to “think” autonomously necessarily gives them the capacity to act in ways that may be contrary to the “rules” they are given. That conclusion, of course, lies at the heart of the academic debate over the wisdom of building artificially intelligent machines in the first place and is the animating force behind many science fiction classics.

*124 Assuming that an inevitable byproduct of giving a machine the capacity to “think” carries with it a risk that the machine will break the rules, how would the law apply to HAL? Suppose Dave returns to Earth and demands compensation for the wrongful deaths of Frank and the three other crewmembers HAL murdered? Unless the law is willing to invest an autonomous thinking machine like HAL with legal personhood, he is beyond the reach of the law. As one court observed, “robots cannot be sued,” even though “they can cause devastating damage.”

The introduction of highly intelligent, autonomous machines may prompt reconsideration of that rule. After all, there is no a priori reason why truly autonomous machines should not be accorded some formal legal status, making them, like corporations and certain trusts, “persons” in the eyes of the law and thus subject to suit. Perhaps justice would be served to put HAL in the dock. Until then, the question will be who, if anyone, should be held to account? Would it be fair to hold liable the companies that designed, programmed, or manufactured HAL 9000, even though they embedded in HAL's “thinking” systems the first rule of autonomous machines--i.e., never harm a human--and even

*125
though the evidence strongly suggests that HAL “taught” himself to defy their instructions? Or should the creators of machines that have the capacity to “think” be held strictly liable whenever anything goes wrong? If so, on what theory? The theory that the wrongful conduct itself is proof of a defect? Or on an insurance-based theory that the creators are in a better economic position to absorb the cost of the injury than the person harmed?

As fanciful as it may seem, these are by no means idle questions. Machines that operate and make decisions independently will at some point be introduced into the market. The first truly autonomous artificial intelligence devices that may test the adequacy of current liability rules may be cars designed to be driver-less, or at least to give human drivers the option to let the car drive itself. Google's driver-less cars have been test-driven on public roads for several years and have logged hundreds of thousands of miles. Although there is a human sitting in the “driver's seat,” the human is, for the most part, playing the role of potted plant. The car is driven by systems that use an array of radar and laser sensors, cameras, global positioning devices, many complex analytical programs and algorithms (and other devices far beyond my comprehension) to actually drive the car in much the same way that humans do, only better. The car “watches” the road, constantly looks out for other cars, pedestrians, obstructions, detours, and so forth, and adjusts its speed and course to account for traffic, weather, and every other factor that affects the safe operation of the vehicle. And they are programmed to avoid a collision with a pedestrian, another vehicle, or an obstacle. Google has claimed that its driver-less cars have yet to have an accident (except when one of its cars was rear-ended)—an impressive record.

It is hardly surprising that Google's experience has thus far validated the hypothesis that autonomous cars would perform far more safely than cars driven by humans. Driving is risky because drivers are humans. Humans inject risks that can be eliminated, or at least considerably mitigated, by intelligent machines. After all, the Google system that drives cars “sees” everything in the vicinity; reacts at speeds humans cannot match; and constantly checks the performance of every component in the vehicle to ensure that it is functioning properly. What's more, the system never gets drowsy or falls asleep, does not drive drunk, does not eat Big Macs or drink hot coffee, does not get distracted, does not talk on the phone or text, and does not get road rage. The system focuses solely on the one activity it is designed for: Driving from point A to point B as safely and efficiently as possible.

Like human drivers, the machines that drive Google cars will on occasion encounter unexpected events that call for snap judgments: a child darting in front of a car; a tree limb crashing down just a few yards ahead; a car running a red light; or a patch of black ice that is undetectable on a moonless night. There are liability rules that come into play when, as a consequence of any of these unexpected events, there is injury to humans or damage to property. No matter how well-designed and programmed self-driving cars are, factors beyond the machine's control virtually guarantee that at some point the car will have an accident that will cause injury of some kind, and will act in ways that are not necessarily ordained by their programming. The question I have been asked to address in this symposium is: What liability rules should society adopt to govern artificially intelligent machines, such as the driver-less cars that soon may be right in front of us on the highway?

Answering this question requires that two different issues be addressed. The first is how to apply the law of products liability on the assumption that any liability concern with the machine is the result of human (but not driver) error—that is, a design or manufacturing defect, an information defect, or a failure to instruct humans on the safe and appropriate use of the product. In my view, the application of these reasonably settled principles is a straightforward one, and there is no justification for treating even autonomous thinking machines differently than any other machine or tool a human may use, except, perhaps, holding them to a higher standard of care.
The second question comes into play if, and only if, fully autonomous machines cause injury in ways wholly untraceable and unattributable to the hand of man. This question, in my view, crystalizes the HAL problem. It is fair to assume that, if driver-less cars become the norm, there will be accidents, perhaps few and far between, that cannot fairly be attributed to a design, manufacturing, or programming defect, and where even an inference of defect may be hard to justify. What should be the rule at that point, especially where the car acts in a way that is at odds with the instructions of its creators? Tort law is ordinarily unwilling to let people injured through no fault of their own bear costs imposed by others. So the question then becomes, “Who pays?” The only feasible approach, it would seem, would be to infer a defect of some kind on the theory that the accident itself is proof of defect, even if there is compelling evidence that cuts against a defect theory. There is precedent for courts making such an inference, which is simply a restatement of res ipsa loquitor. If that is the right choice to make (and I argue it is), then there is the secondary question of how, if at all, should the law apportion liability among designers, programmers, manufacturers, and others involved in the vehicle's creation? Or, as suggested above, should liability simply be assigned to the vehicle itself?

Of course, the solution suggested by existing law would be to hold the vehicle's manufacturer liable and let the manufacturer seek indemnity or contribution from other potentially responsible parties, if any. But that approach may be nothing more than an empty gesture. If it is in fact impossible to identify the cause of the accident, then the manufacturer would likely have no reasonable grounds for an indemnity or contribution action, and would thus be saddled with the entire judgment. That result might make sense if the manufacturer is in the best position to bear the loss. Otherwise, it might be fairer to apportion responsibility among all of the parties that participated in building and maintaining the vehicle's autonomous systems, on the ground that the cost of error is better spread among all potentially responsible parties or among the parties who could more efficiently guard or insure against the loss. The other approach would be to hold the vehicle itself responsible, assuming, of course, that the law is willing to confer legal “personhood” on the vehicle and require the vehicle to obtain adequate insurance.

I. AUTONOMOUS MACHINES AND LIABILITY RULES RELATING TO HUMAN (BUT NOT DRIVER) ERROR

The key question in discussing liability is who bears the costs when something goes wrong and injury occurs. Driver-less cars, at least those that might be on the road within the next decade, will reduce but will not entirely eliminate the possibility of driver error. As I understand it, the system Google is developing permits the person sitting in the driver’s seat to take control of the car whenever he or she wants. At this point, however, it is unclear whether Google intends for the driver to have control at some points: For instance, while driving through high pedestrian traffic residential areas, construction zones, or during especially inclement weather, even though Google's driver-less cars have performed well in those conditions. But the Google system also has a voice command that instructs the driver to take control of the vehicle, presumably when the control system detects some kind of error or is unable to determine what it should do. The point of this essay is to explore liability concerns where there is no possibility that a human driver caused or contributed to the accident. So what follows is based on the understanding that the person in the driver's seat is disengaged from any aspect of driving the car, and that the car is under the sole control of the car itself. Simply taking human drivers out of the equation should mean, in the long run, a considerable drop in accidents.

But even in the best of circumstances, unexpected and improbable events occur, things go wrong, and someone or something is injured. Product liability law provides a framework for resolving claims by parties injured when things go wrong. With a “product” like an autonomous car, the law groups those possible failures into familiar categories: design defects, manufacturing defects, information defects, and failures to instruct on appropriate uses.
Before examining these theories of liability, it is useful to pause to consider whether the standard of care to be applied to driver-less cars will be different than the standard applied to cars driven by humans. There is every reason to think that the answer will be “yes,” and that fact may bear on the analysis that follows.

As unlikely as this may seem, the standard of care question was addressed by a court more than a half-century ago. In Arnold v. Reuther, an intermediate appellate court in Louisiana questioned whether autonomous cars will be held to a higher standard of care than cars driven by humans. Arnold involved a negligence claim by Mrs. Arnold against Mr. Reuther, whose car struck Mrs. Arnold when she darted across Canal Street in New Orleans as he was making a left-hand turn. Although the facts were controverted, the court summarized the relevant, undisputed facts this way: “[J]ust before the car of Reuther reached a point opposite that at which Mrs. Arnold stood on the sidewalk among the other pedestrians, she, at a hurried pace, entered the street and attempted to cross;” the other pedestrians did not attempt to cross. According to the court, Mr. Reuther testified that “he did not see Mrs. Arnold as she left the sidewalk and, in fact, did not see her until she was actually ‘in front’ of his car.”

Mrs. Arnold did not dispute that she attempted to cross the street without looking and that her conduct was negligent. Her argument was that Mr. Reuther had the “last clear chance” to avoid the accident, but failed to do so and should be held liable for that reason. The court disagreed, concluding that Mr. Reuther's efforts, although ineffective, were sufficient to avoid liability because, after all, he was only human. The court said:

A human being, no matter how efficient, is not a mechanical robot and does not possess the ability of a radar machine to discover danger before it becomes manifest. Some allowances, however slight, must be made for human frailties and for reaction, and if any allowance whatever is made for the fact that a human being must require a fraction of a second for reaction and cannot respond with the mechanical speed and accuracy such as is found in modern mechanical devices, it must be realized that there was nothing that Reuther, a human being, could have done to have avoided the unfortunate result which the negligence of Mrs. Arnold brought upon herself.

Fifty years later, the court's insight seems remarkably prescient. Had Mr. Reuther been sitting behind the wheel of a driver-less car, it is quite possible that, as the Louisiana court forecast, the accident would have been averted; the car would have immediately detected Mrs. Arnold's ill-considered dash across the street and rapidly applied the brakes (after confirming that there was no vehicle behind it that would be unable to stop safely) to come to a stop before striking Mrs. Arnold and before Mr. Reuther was even alert to the danger.

As this example illustrates, in all likelihood, in the first accident case involving a driver-less car, the court will not ask whether the self-driving car performed as well as a reasonable human. The question in Arnold was one of negligence: Did Mr. Reuther behave reasonably under the circumstances? And the court said he had. But with a driver-less car, and no human at the controls, the focus will be on whether the car performed as well as it should have. To be sure, one could state that question in the language of negligence, and there may be questions about whether the car or some of its key components were designed, manufactured, or maintained negligently. But the court in the first driver-less car case will likely ask whether the car involved in the accident performed up to the standards achievable by the majority of other driver-less cars, as well as the performance specification set by the car's manufacturer.
Just how demanding would that standard be? Under contemporary products liability law, the standard would likely be set by a risk-utility analysis of the sort described in the Restatement (Third) of Torts. To be sure, in the past, cases like Arnold proceeded under negligence theories, in large part because the focus was on the conduct of humans, not the performance of machines. But since then, as negligence claims involving product failures became more difficult to prove, strict liability principles took root to govern product liability cases. Products liability plaintiffs often plead both negligence and strict liability claims. But cases involving complex products are now typically viewed through the lens of products liability law. Take Mrs. Arnold's case. Imagine that the car that struck her was a driver-less car, and taking her cue from the court, she claims that the car Mr. Reuther was operating (but not driving) should have been designed in a way that would have averted the accident. That is a question of design, technological and economic feasibility, and consumer expectations, and there is no doubt that there will be design defect cases brought against the manufacturers and designers of self-driving cars. Under current law, Mrs. Arnold would have a design defect claim against the manufacturer, but whether she would have a winning case is far from clear, and would depend heavily on the facts leading up to her accident.

To explain why that is so, consider two scenarios, both of which are based on the supposition that there are no mechanical or manufacturing defects with the car. In scenario one, assume that the evidence shows that a driver-less car's sensors were capable of detecting Mrs. Arnold moving into the car's path at a point when she was 15 feet away. And suppose further that, to the extent that there is an industry “collision avoidance” standard, it is that a driver-less car should be able to detect objects (human or not) in its path and avoid a collision (by applying the brakes or by taking evasive maneuvers) if the object is detected within 12 feet of the car. In scenario two, we reverse the numbers so now the industry “collision avoidance” standard is 15 feet, but it is clear that Mrs. Arnold was not detected by the sensors until she was 12 feet from the car.

Determining the proper test for design defects has been the central and most contentious question in products liability law for the past half-century, and it does not appear that the debate has been conclusively resolved. Under current law, there are two potentially viable liability theories that plaintiffs may be able to employ to demonstrate that a product was defectively designed. In the jurisdictions that still follow the Restatement (Second) of Torts or some variant of it, Mrs. Arnold could proceed under a “consumer expectations” test. Putting aside that courts and academics have been critical of using a consumer expectations test in design defects, and especially in cases involving complex products, that claim permits juries to find for the plaintiff whenever a product is in a “defective condition unreasonably dangerous to [a] user or consumer.” An “unreasonably dangerous” defect is one that makes a product “dangerous to an extent beyond that which would be contemplated by the ordinary consumer who purchases it, with the ordinary knowledge common to the community as to its characteristics.” Although there are difficulties with this test (discussed below), it does provide Mrs. Arnold a path to get her case before a jury, where jurors can then “draw[] their own reasonable conclusions as to the expectations of the ordinary consumer and the knowledge common in the community at large.”

The second test is the risk-utility test laid out in the Restatement (Third) of Torts: Products Liability, which would require Mrs. Arnold to show that “the foreseeable risks of harm posed by the product could have been reduced by the adoption of a reasonable alternative design.” That is a formidable burden. The word “reasonable” is intended to import a quantitative cost-benefit analysis into the test, and thus under the Restatement's formulation, Mrs. Arnold would have to shoulder the burden of proving that some modified version of the product would have avoided the collision at a cost that is reasonable in relation to the degree of harm reduced.
Under scenario one, where the car's sensors should have detected Mrs. Arnold's presence in time to avoid the collision but failed to do so, Mrs. Arnold is likely to prevail under each theory. She would likely prevail under the consumer expectations test because consumers would reasonably expect that Mr. Reuther's driver-less car would meet or exceed the existing 12-foot “collision avoidance” industry standard. The analysis conducted under the risk-utility test should come out in her favor as well. She would not have to advocate in favor of an “alternative” design; her argument would be that the market had already weighed the relative costs and benefits of the 12-foot collision avoidance standard and determined that the cost of such a standard is justified by the corresponding risk reduction.

Scenario two raises much more difficult questions because most, perhaps all, driver-less vehicles would have been unable to stop or take evasive action quickly enough to avoid the accident. For starters, it is not clear that Mrs. Arnold could prevail under a consumer expectations test. To be sure, given the indeterminacy of the test--the standard is, after all, what reasonable consumers say it is--it could be her salvation or her undoing. On the one hand, the consumer expectations test might work to Mrs. Arnold's advantage. There has been considerable hype about the enhanced safety of driver-less cars--so much so that some experts forecast that these cars, once we rid the roads of drivers, will put an end to accidents. Reacting to that hype, consumers may expect (putting aside questions of reasonableness) driver-less cars to be designed in a way that would avert accidents even where, as is the case with Mrs. Arnold, the initial “fault” lies with the person and the machine had at best a second or two to detect and react to the impending accident. Indeed, even if the industry's evidence shows that the 15-foot collision avoidance standard pushes the limits of technological and economic feasibility, jurors might conclude that designers are capable of doing better and hold the manufacturer of Mr. Reuther's car liable on that basis. Whether the court would let a verdict like that stand is open to question, but the consumer expectations test is a double-edged sword; expectations are often shaped by industry, but industry standards do not necessarily define the outer boundary of enforceable consumer expectations.

On the other hand, one downside of the test is that consumer expectations are generally a reflection of the existing state of the marketplace. Manufacturers, through advertising and other communications with consumers, play a key role in shaping consumer expectations. Unless the manufacturer makes inflated and unjustified representations about its product's performance, consumers are likely to expect that their products will perform in a way that is consistent with prevailing standards as articulated by the products' manufacturers, even if better and safer products are achievable at a nominal cost. Under this theory, it may be fair to assume that consumers would expect driver-less cars to operate in a safe manner and might even have a reasonable expectation that driver-less cars will deliver higher levels of safety than cars driven by humans. But it seems a stretch to assume that consumers would have formed a reasonable expectation that driver-less cars can stop quickly enough to avoid hitting a person who darts in front of the car. This point might take on added force if the manufacturers of driver-less cars seek to dampen consumer expectations about the safety features of the cars, if for no other reason than to quell wholly unrealistic expectations that might lead consumers to think, as Mrs. Arnold may have thought, that no matter how risky her conduct was, the car would be able to stop quickly enough to avoid impact. For these reasons, even under the consumer expectations test, perceived by many to be more plaintiff-friendly, Mrs. Arnold would face a tough and uncertain path to victory.

Equally uncertain is whether the risk-utility test embodied in the Restatement (Third) of Torts would provide Mrs. Arnold a road to recovery. Her burden would be to show that the incremental benefit of altering the car's design to improve stopping distance from 15 feet to 12 feet would be worth the cost--that is, that the value of the accidents avoided would equal or exceed the added cost of the alternative design. Presenting evidence to that effect would be difficult. After all, Mrs. Arnold would need to retain an expert engineer to demonstrate that a reduction in the stopping distance is technologically feasible, as well as another expert (likely an economist) to establish that the savings achieved by the
reduction would outweigh the attendant costs of modifying the vehicle.  

Her experts would also have to do battle with the manufacturer's experts, who would likely claim that the design pushed the limits of technological feasibility.  

To be sure, if she could make the showing that additional modifications could have made the vehicle safer at a reasonable cost, she might have a case. But finding and retaining experts to put on a case might be a dauntingly expensive enterprise for an individual plaintiff to bear.

Mrs. Arnold's path to victory would be far surer if it turned out that there was a manufacturing defect or component failure that caused or contributed to the accident. A product with a manufacturing defect is a product with an unintended flaw; that is, the finished product does not conform to the manufacturer's own specifications or requirements. A plaintiff can prevail in a manufacturing defect case if she can prove that the product does not conform to the manufacturer's specifications and that the defect played a role in causing the accident. For autonomous cars, there are certain to be instances in which vital components of the system fail because of a manufacturing or maintenance defect. For example, assume that one of the key components responsible for keeping track of objects that might enter the path of the car (for instance, a radar or laser sensor) failed on Mr. Reuther's car prior to the accident with Mrs. Arnold. Assuming that the failing component can be identified, and assuming that there is a link to the failed component and the accident, liability can be established and properly allocated. If the component is analyzed and does not meet the manufacturer's specification, then the case is an easy one, and liability may be appropriately meted out.

Even if there is no tangible sign of defect, Mrs. Arnold may have a viable claim, at least in some jurisdictions, if she can prove that the component malfunctioned and, for instance, resulted in the car failing to stop in a manner consistent with the manufacturer's specifications. The law permits such a claim to go forward so long as the component was properly used, properly maintained, and had not been altered or damaged in a way that might have caused the malfunction. But take note: In cases involving other autonomous machines, liability has been difficult to establish where alternative theories of liability are present.

Finally, there is the issue of information defects. Although it is unlikely that Mrs. Arnold would have a claim under this theory, the “duty to train” often plays a role when humans interact with, and have to operate, complex machines. The duty to train may play an important role in the introduction of driver-less cars onto the market. Driver-less cars are highly sophisticated systems, and learning how to “drive” and “operate” them may be difficult for consumers used to being in control while they are in the “driver's seat.” Especially if the “driver” must remain alert and ready to assume the conventional role of “driver,” the burden of teaching consumers how to manage those tasks may fall to manufacturers, and they may face liability if they fail to adequately communicate information about the safe use of driver-less cars.

As this brief review shows, the introduction of truly autonomous vehicles is unlikely to present legal issues that tax our current product liability regime, at least so long as the product failure linked to the injury can be traced to, or reasonably imputed to, the activity of an identifiable person or legal entity. To be sure, the increasing complexity of automotive technology will provide even more ammunition to those who criticize the continued use of a consumer expectation test in design defect litigation. Supporters of the test may have a difficult time defending its use in cases involving driver-less cars, given the vehicles' intricate and interdependent technologies and the argument from opponents that consumers will not have a sufficient basis to form a reasonable judgment about what level of safety to expect. On the other hand, as many academics have suggested, when properly applied, the risk-utility test laid out in the Restatement (Third) of Torts does embody some elements of consumer expectations, and that test may adequately safeguard the interests of those injured by driver-less vehicles.
II. MACHINES WITHOUT PRINCIPALS AND LIABILITY RULES RELATING TO INEXPLICABLE ACCIDENTS

Thus far, this Essay has addressed the application of conventional products liability principles to machines that are capable of operating independently of human direction, but where the machine's failure can fairly be attributed to some act or omission by a human that can be said to have “caused” the accident. As we have seen, apart from the vexing question of whether the “operator” or manufacturer of the driver-less car should be the defendant in a case where the car and not the driver had control, products liability law will apply in much the same way as it would apply if the car involved in the accident were driven by a human. And as we have also seen, resolution of the proper defendant question is likely to have little bearing on the important question of “Who pays?” In a fault-based system, drivers will bear the loss when they are responsible for the accident, and manufacturers will bear the loss when the driver-less car they have designed or manufactured fails. But in all of these scenarios, it is easy to avoid the question of agency because some human--Mrs. Arnold, Mr. Reuther, the designer, the programmer, or the manufacturer of the vehicle or its component parts--was the key causal link in the events that led to the injury and thus can be assigned the blame.

Now switch gears and assume that an accident occurs involving a truly autonomous vehicle where it is impossible to reasonably assign the responsibility for the accident to a human. The hard question that may arise when autonomous machines are involved in accidents is what should the liability rule be when it is unreasonable to infer that the accident was caused by a design or manufacturing defect. It may be that those cases prove to be the null set. After all, an inference of defect is reasonably drawn when a product fails, even when a defect cannot be determined by engineers, when the failure occurs with some frequency and the failure follows a common pattern. In those circumstances, courts routinely apply principles of res ipsa loquitor and conclude that the car, not the driver, is at fault. 81

As one example, consider the recent flood of design defect cases brought against Toyota alleging that certain Toyota-manufactured vehicles (both Toyota and Lexus) are prone to “sudden acceleration” through no act by the driver and the acceleration cannot be stopped by the driver, by either disengaging the accelerator pedal or by applying the brakes. Notwithstanding an exhaustive inquiry, engineers have been unable to identify a specific design or manufacturing defect that causes the uncontrolled acceleration, although theories abound. 82 Nonetheless, because of the substantial number and common features of these complaints, Toyota has decided not to contest liability. It settled a class action lawsuit for economic loss due to the defect brought by owners of affected Toyota and Lexus vehicles for $1.3 billion and other relief. 83 More recently, Toyota took steps to settle the pending 400 personal injury cases against it after an Oklahoma jury, applying the doctrine of res ipsa loquitor, awarded the plaintiffs $3 million. The jury apparently concluded that, even though the plaintiffs could not isolate the cause of sudden acceleration, the accident was more likely caused by the car than the driver. 84 As the Toyota case makes plain, existing products liability law is well-positioned to address cases where the evidence strongly suggests a defect, but technology cannot isolate the cause.

What about failures involving driver-less vehicles that are rare, untraceable to any defect, and inexplicable in that the vehicle's actions are incompatible with the vehicle's decision-making design? Who, if anyone, should bear liability then? To make the question more concrete, suppose daredevil Mrs. Arnold again darts in front of a driver-less vehicle. In this scenario, assume that the vehicle's designers have programmed instructions that it should at all costs avoid striking a human. And further assume that, in order to avoid hitting Mrs. Arnold, the vehicle has only two choices: It can either (1) take evasive action by making a sharp turn that would likely avoid any impact with Mrs. Arnold, but would also likely result in the vehicle striking a brick wall (thereby risking the safety of the vehicle's occupants), or (2) apply the brakes with great force, even though, given the distance and vehicle's speed, that course of action would not guarantee
that the vehicle would stop before hitting Mrs. Arnold; nor would it ensure that there would be no collision between it and the car following closely behind. 86

Suppose that the vehicle does not adhere to its basic instruction to avoid a collision with a human if at all possible and takes neither of these courses of action. Instead, the car applies its brakes gently, decreasing the car's speed, and avoiding the brick wall and a possible collision with the car behind it, but nonetheless striking and injuring Mrs. Arnold. As noted, the car was not “supposed” to make the choice it did. The engineers and programmers who designed the car's autonomous driving system “instructed” the car to avoid the possible collision with Mrs. Arnold and the car behind it by braking and by steering itself (and its passengers) into the brick wall, recognizing that the vehicle was sufficiently crash-worthy to sustain the impact without unreasonable risk to the safety of the car's occupants.

In some sense, this is the HAL question presented in a more benign form. It may be that truly intelligent machines will learn to adapt the instructions they initially receive from humans to circumstances not directly forecast at the time of their creation. And perhaps these machines will learn to internalize values that are not the ones their creators tried to embed. HAL, for instance, was not programmed to value self-preservation, but we know that he held that value dearly, and placed it above human life. 87

*145 The intended lesson of this hypothetical is that at some point it will be hard to conceptualize truly intelligent machines as mere agents or tools of humans. Agency principles can take us only so far. A machine that can define its own path, make its own decisions, and set its own priorities may become something other than an agent. Exactly what that may be, though, is not a question that the law is prepared to answer. 88

Nor does current product liability law yield answers that are particularly satisfying. Consider the hypothetical sketched out above. In that case, even though the car might have been capable of avoiding Mrs. Arnold, her recovery would be uncertain, even if her claim were stronger. First, we can rule out a manufacturing defect claim; the mechanical features of the car functioned perfectly. To the extent that Mrs. Arnold has a path to recovery, it lies with a design defect claim. The consumer expectations test, even in those jurisdictions that have crafted a hybrid form of consumer expectations/risk-utility test, is unlikely to aid Mrs. Arnold. Whatever else this case may be, it is not an “ordinary” case where it is fair to presume that jurors will have reasoned expectations about how the vehicle should have performed.

Nor will Mrs. Arnold get much traction under the risk-utility theory, which, in a case like our hypothetical, would likely place the burden of showing a design alternative on the plaintiff. Again, the complexity and sophistication of driver-less cars, and the complications that will come with the fact patterns that are likely to arise, are going to make proof of wrongdoing in any individual case extremely difficult. The question whether there are alternative designs that would make driver-less cars even safer and prevent one-off accidents will be so infused with complex technical and economic questions that individual cases will be difficult and expensive to try, and even harder for jurors to resolve.

This case illustrates the point. Even though the vehicle did not perform in a way that was consistent with its designer's intentions, the vehicle did make an objectively reasonable choice: The vehicle took swift and decisive action to minimize the harm to Mrs. Arnold while safeguarding the vehicle's passenger from possible harm. The jurors may see that choice as a sensible one. And we know that Mrs. Arnold cannot argue that the better design choice is a return to the status quo ante, where cars were driven by humans. Even with today's safety features--automatic braking, anti-lock brakes, traction control, airbags, 89 backup cameras, “blind spot” cameras, and the like --the rate of accidents and injuries on the roads in the United States remains frighteningly high. 90 Whatever flaws may emerge in driver-less cars, their introduction will almost certainly make a considerable dent in the accident and injury rate in the United States.
In my view, in cases where driver-less cars fail and cause injuries to persons or property and it would be unreasonable to attribute the failure to the vehicle's manufacture or design, the law will need to fashion a response that best serves the collective interests of the affected parties. My proposal is to construct a system of strict liability, completely uncoupled from notions of fault for this select group of cases. A strict liability regime cannot be based here on the argument that the vehicles are “ultra-hazardous” or “unreasonably risky” for the simple reason that driver-less vehicles are likely to be far less hazardous or risky than the products they replace. Indeed, it is precisely because these machines are so technologically advanced that we expect them not to fail. For these reasons, a true strict liability regime will be needed; one that does not resort to a risk-utility test or the re-institution of a negligence standard for the simple fact that those tests will be difficult, if not impossible, for the injured party to overcome.

In this instance, a system of strict liability would, in effect, impose a court-compelled insurance regime to address the inadequacy of tort law to resolve questions of liability that may push beyond the frontiers of science and technology. There are four strong policy reasons to establish a strict liability regime for this category of cases.

First, providing redress for persons injured through no fault of their own is an important value in its own right. The idea that individuals should bear a loss that is visited upon them, even when the causal failure is inexplicable, runs counter to basic notions of fairness, compensatory justice, and the apportionment of risk in society.

Second, a strict liability regime is warranted because, in contrast to the injured party, the vehicle's creators are in a position to either absorb the costs, or through pricing decisions, to spread the burden of loss widely. After all, it is not unreasonable that the costs of inexplicable accidents be borne, at least in part, by those who benefit from risk-reducing, innovative products.

Third, a strict liability regime will spare all concerned the enormous transaction costs that would be expended if parties had to litigate liability issues involving driver-less cars where fault cannot be established. The first law of litigation is as follows: As the complexity of products rises geometrically, the cost of litigating products liability cases increases exponentially. It is better to spend money compensating the injured than paying lawyers and experts. Indeed, although Toyota has not discussed this issue publicly, it is possible that Toyota decided to settle the sudden acceleration cases not because it feared losing them in court, but because it feared that litigating each case to judgment, even if it won, would cost the company far more than settling the cases because of the complexity of the cases. If so, Toyota would not be the first company to settle a case to avoid transaction costs, regardless of the case's merits.

And fourth, a predictable liability regime may better spur innovation than a less predictable system that depends on a quixotic search for, and then assignment of, fault. If driver-less cars deliver even a fraction of their promised benefits of less carnage on the highways, more leisure time for consumers, and reduced fuel consumption, they will provide a substantial benefit to society. A liability system should not stifle innovation; it should encourage responsible innovation. Stability, coupled with a cost-spreading approach, would doubtlessly serve that goal better than an uncertain fault-based liability system.

Assuming that the courts can be persuaded to adopt a strict liability regime, the question then becomes, “Who bears the cost?” Should the burden be placed on the operator, owner, the manufacturer, the programmers, the designers, or all of them, or should the law simply declare that the driver-less car itself is a legal “person” and require it to insure itself against liability? Under prevailing products liability law, if someone is injured by the failure of an automobile (and not the driver), the manufacturer is the target of choice. It is the manufacturer who bears front line responsibility for design and manufacturing defects, and is generally the principal if not the only defendant in litigation. That allocation
of responsibility makes sense, because the manufacturer sets the price for the vehicle, and so the manufacturer can build in an “insurance premium” into the vehicle's sale price to offset expected liability costs.  

There are at least two concerns about making the manufacturer shoulder the costs alone. One is that with driver-less cars, it may be that the most technologically complex parts--the automated driving systems, the radar and laser sensors that guide them, and the computers that make the decisions--are prone to undetectable failure. But those components may not be made by the manufacturer. From a cost-spreading standpoint, it is far from clear that the manufacturer should absorb the costs when parts and computer code supplied by other companies may be the root cause. Second, to the extent that it makes sense to provide incentives for the producers of the components of driver-less cars to continue to innovate and improve their products, insulating them from cost-sharing even in these kinds, of one-off incidents seems problematic. The counter-argument would of course be that under current law the injured parties are unlikely to have any claim against the component producers, and the manufacturer almost certainly could not bring an action for contribution or indemnity against a component manufacturer without evidence that a design or manufacturing defect in the component was at fault. So unless the courts address this issue in fashioning a strict liability regime, the manufacturer, and the manufacturer alone, is likely to bear all of the liability.  

*149* One way--perhaps the only way under conventional law--to untie this Gordian Knot would be to apply a variation on the settled doctrine of “common enterprise” liability. Under that theory, “each entity within a set of interrelated companies may be held jointly and severally liable for the actions of other entities that are part of the group.” The variations here would require a substantial deviation from the doctrine. For one thing, the proposed liability theory would not require that the companies function jointly; all that would be required is that they work to a common end--to design, program, and manufacture a driver-less car and its various component parts. For another, there is the question of fault. The common enterprise doctrine is ordinarily applied to ensure that once liability has been established, all of the participating wrongdoers are held to account. Here, of course, there would be no “wrongdoers.” There would instead be an inference of liability drawn by operation of law to protect a blameless party (the person who sustained injured) by making others bear the cost. But the basic point is the same: A common enterprise theory permits the law to impose joint liability without having to lay bare and grapple with the details of assigning every aspect of wrongdoing to one party or another; it is enough that in pursuit of a common aim the parties engaged in wrongdoing. That principle could be engrafted onto a new, strict liability regime to address the harms that may be visited on humans by intelligent autonomous machines when it is impossible or impracticable to assign fault to a specific person.  

There is one other possible solution that avoids the legal fiction of having the creators of the vehicle bear the cost even without a hint of *150* fault, and that is to have the vehicle itself bear the cost. Instead of suing the manufacturer, let the injured party do what is now not possible--sue the vehicle. As suggested earlier, at some point the courts will have to wrestle with the underlying question of how to treat machines that are agents of no one--i.e., machines without principals. One solution would be to reconceptualize these autonomous, intelligent machines as entities with the status of a “person” under the law. Conferring “personhood” on these machines would resolve the agency question; the machines become principals in their own right, and along with new legal status would come new legal burdens, including the burden of self-insurance. This is a different form of cost-spreading than focusing on the vehicle's creators, and it may have the virtue of necessitating that a broader audience--including the vehicle's owner--participate in funding the insurance pool, and that too may be more fair.  

CONCLUSION
The introduction of highly sophisticated autonomous machines may be literally around the corner. Truly autonomous machines may be driving cars through our neighborhoods or piloting drones that fly above our heads sooner than we think. So long as we can conceive of these machines as “agents” of some legal person (individual or virtual), our current system of products liability will be able to address the legal issues surrounding their introduction without significant modification. But the law is not necessarily equipped to address the legal issues that will start to arise when the inevitable occurs and these machines cause injury, but when there is no “principal” directing the actions of the machine. How the law chooses to treat machines without principals will be the central legal question that accompanies the introduction of truly autonomous machines, and at some point, the law will need to have an answer to that question.

Footnotes


2 Id. at 413.


5 See Trafton, supra note 3 (reporting that “Domo” is based in part on “Rosie”).


See infra notes 15-16 and accompanying text.

See 2001: A Space Odyssey, supra note 9; 2001: A Space Odyssey, Quotes, IMDb, http://www.imdb.com/title/tt0062622/quotes (last visited Feb. 26, 2014) (HAL's complete answer to the interviewer's question is as follows: “The 9000 series is the most reliable computer ever made. No 9000 computer has ever made a mistake or distorted information. We are all, by any practical definition of the words, foolproof and incapable of error.”).

Id. (Dave's full response to the question is as follows: “Well, he acts like he has genuine emotions. Um, of course he's programmed that way to make it easier for us to talk to him. But as to whether he has real feelings is something I don't think anyone can truthfully answer.”).

See 2001: A Space Odyssey, supra note 9; 2001: A Space Odyssey, Quotes, IMDb, http://www.imdb.com/title/tt0062622/quotes (last visited Feb. 26, 2014) (“I'm afraid. I'm afraid, Dave. Dave, my mind is going. I can feel it. I can feel it. My mind is going. There is no question about it. I can feel it. I can feel it. I'm a...fraid.”). The same intelligent machine-versus-humans theme was explored with far less nuance but much more mayhem in the 1984 film The Terminator, which tells the story, set in 2029, of how an artificially intelligent defense network, “Skynet,” becomes self-aware and decides to wipe-out mankind through a nuclear holocaust. Skynet nearly succeeds. Only a few humans survive, and they form a resistance movement. Over time, the resistance gains enough strength that it is on the verge of beating the machines. To regain the offensive, Skynet sends the “Terminator,” played by Arnold Schwarzenegger (people forget that he was the villain in the first film), back to 1984 to kill Sarah Connor, then pregnant with John Connor, who grows up to lead the resistance against Skynet. Much entertaining violence ensues, and the film ends inclusively. See The Terminator (Orion Pictures Corp. 1984). John Connor and his mother survive, in part because the resistance sends its own highly intelligent robot/warrior back in time to defend them, but there is no doubt that they will face another onslaught from Skynet terminators. Id. These onslaughts come in an unending number of sequels involving ever more intelligent and dangerous robotic assassins, but John Connor nonetheless eludes them all, mainly because of help from a highly skilled robot warrior sent by the resistance to protect him (played by Arnold Schwarzenegger). See, e.g., Terminator 2: Judgment Day (TriStar Pictures 1991); Terminator 3: The Rise of the Machines (Warner Bros. Pictures, Columbia Pictures 2003). But the film's basic point is a warning to humans: Be careful what we wish for. Artificially intelligent machines that have the capacity to “think” for themselves should not be trusted because they may turn on humans.

See infra notes 15-16 and accompanying text.

There has been extensive litigation over the safety of surgical robots, especially the “da Vinci” robot. Not surprisingly, no one has claimed that the robot itself bears any liability; the claims all proceed on some form of agency theory. See, e.g., O'Brien v. Intuitive Surgical, Inc., No. 10 C 3005, 2011 WL 304079, at *1 (N.D. Ill. Jul. 25, 2011) (granting summary judgment to robot's manufacturer); Mracek v. Bryn Mawr Hosp., 610 F. Supp. 2d 401, 402 (E.D. Pa. 2009) (same), aff'd, 363 F. App'x 925 (3d Cir. 2010).

It is useful to contrast the complexity of driver-less cars to what we think of as highly sophisticated machines in widespread use today. For instance, most commercial airplanes have “auto pilots,” which shift control of the aircraft to a computer that “flies” the plane. So too do most vessels. But in comparison to driver-less cars, the autopilot devices perform a relatively simple set of tasks. Autopilots keep the plane or vessel on a course determined by the pilots, and do so by controlling for minor variations in winds and currents, but generally without reference to other traffic. For that reason, pilots have a duty to remain vigilant--while the machine may have the controls, the pilots are responsible for monitoring other traffic and ensuring that the autopilot is working correctly. See, e.g., In re Korean Air Lines Disaster of Sept. 1, 1983, 932 F.2d 1475, 1478 (D.C. Cir. 1991); Boucvalt v. Sea-Trac Offshore Servs., Inc., 06-103 (La. App. 5 Cir. 10/17/06); 943 So. 2d 1204, 1208 . In contrast, for autonomous cars, one key goal is to reduce the oversight required by the driver, both to maximize safety and to capture the possibility of increased productivity during long commutes. The systems that control autonomous cars are required to navigate through complex and rapidly changing environments--e.g., traffic, weather, detours and the like--and are responsible for making critical decisions--e.g., what route to take, what lane to be in, what exit to take, and so forth. See Dylan LeValley, Note, Autonomous Vehicle Liability--Application of Common Carrier Liability, 36 Seattle U. L. Rev. 5, 7 (2013). And there may be good reasons why the designers of driver-less cars would want to discourage human drivers from intervening. A driver
intervening to prevent what he perceives to be a possible accident may be less equipped to handle the situation than the car's autonomous driving system; he may not have monitored blind spots, heard the siren of an approaching emergency vehicle, or been privy to a communication sent by another vehicle in close proximity warning of some danger, and his reaction time is certainly going to be slower than the car's. See id. at 16 (pointing out that Air France flight 447 crashed over the Atlantic because of pilot error. When the plane's autopilot sensors stopped working due to the build-up of ice, the pilots took over control of the plane, and reacted in a way that exacerbated the icing problem, causing the plane to crash).

17 Ugo Pagallo, The Law of Robots: Crimes, Contracts and Torts 2 (2013). This Essay draws significantly on Professor Pagallo's path-breaking work, which broadly and ambitiously synthesizes the technical, philosophical, and legal questions surrounding the advent of fully autonomous, artificially intelligent machines.

18 Nick Bostrom, When Machines Outsmart Humans, 35 Futures 759, 763 (2003). Artificial intelligence theorists use the term “singularity” or “technical singularity” to describe the moment in time, purely hypothetical at this point, when machines exceed human intelligence. At that point, these theorists argue, machines will become fully sentient, and they will pose a raft of complex philosophical and legal questions with which society will have to wrestle. See, e.g., Nick Bostrom & Milan Cirkovic, Global Catastrophic Risks (2008); Ray Kurzweil, The Singularity is Near: When Humans Transcend Biology 135-36 (2005); Hans Moravec, Robot: Mere Machine to Transcendent Mind 61 (1999). Ray Kurzweil predicts that “singularity” is now within reach, and will be achieved within fifteen years. See Ben Rossington, Robots 'Smarter Than Humans Within 15 Years,' Predicts Google's Artificial Intelligence Chief, Mirror News (Feb. 2, 2014), http://www.mirror.co.uk/news/technology-science/technology/ray-kurzweil-robots-smarter-humans-3178027. This Essay does not suggest that the first generation of fully autonomous, artificially intelligent machines will necessarily meet the various definitions of “singularity” that have been propounded. They probably will not. For our purposes, it is sufficient to focus on machines that are capable of “learning” to the point where they can take actions in ways that are not necessarily pre-ordained by the programs that enable them, and thus present the difficult agency problems discussed in Part II, infra.

19 See generally Restatement (Third) of Agency §7.07 (2006) (“An employee acts within the scope of employment when performing work assigned by the employer or engaging in a course of conduct subject to the employer's control. An employee's act is not within the scope of employment when it occurs within an independent course of conduct not intended by the employee to serve any purpose of the employer.”); Lev v. Beverly Enters.-Mass., Inc., 929 N.E.2d 303, 308 (Mass. 2010).

20 The most famous exposition of the “law” of robots comes from Isaac Asimov's I, Robot, where he lays out the Three Laws of Robotics: (1) a robot may not injure a human being, or, through inaction, allow a human being to come to harm; (2) a robot must obey the orders given it by humans, except where such orders would conflict with the First Law; (3) a robot must protect its own existence, as long as such protection does not conflict with the First or Second Law. Isaac Asimov, I, Robot 37 (1950). See generally Arthur C. Clarke, 2001: A Space Odyssey (1968).

21 Once an autonomous machine decides for itself what course of action it should take, the agency relationship becomes frayed or breaks altogether. See Restatement (Third) of Agency §7.07 (2006); id. §7.03 (describing that a principal is subject to vicarious liability for an agent's actions only when the agent is acting within the scope of employment).


MACHINES WITHOUT PRINCIPALS: LIABILITY RULES AND..., 89 Wash. L. Rev. 117

24 See Bostrom, supra note 22, at 77. As noted above, for our purposes, it is not essential that the machine has the capacity to actually choose to break a “rule”; it is enough that the machine's programming does not necessarily determine how the machine will act in all situations, leaving the machine to “learn” how to make decisions when confronted with a situation not within the contemplation of the machine's programmers.

25 Here we put aside any question of criminal liability. At some point, though, the science of artificial intelligence may evolve to the point where concepts of criminality will have to be applied to highly intelligent, autonomous machines. And perhaps that time will come sooner than many think. That was a central message of the 1982 science-fiction thriller, The Blade Runner (Warner Bros. 1982), which was set in 2019, and involved a detective played by Harrison Ford chasing down “replicants”—life-like robots—who were, like HAL, killing humans. Id.; see also Blade Runner: 30th Anniversary, http://bladerunnerthemovie.warnerbros.com/ (last visited Feb. 12, 2014).

26 United States v. Athlone Indus., Inc., 746 F.2d 977, 979 (3d Cir. 1984) (discussing how the manufacturer of a defective robotic pitching machine is liable for civil penalties for the machine's defects). Of course, the court's quip was not completely accurate. The court's point was that, because the robot lacked legal capacity, the robot could not be sued in personam. But the court's conclusion necessarily means that an in rem or quasi in rem action against a robot would be maintainable, precisely because the robot was an “object” rather than a legal “person.”

27 As discussed in more depth infra, this proposal is not as far-fetched as it may seem, and may not require an upheaval in current law to achieve. Conceptualize a highly intelligent machine as one might think of a corporation, which through the income it earned or its intrinsic value was capitalized. There is no reason why the machine itself could not bear liability in case of wrongdoing. For driver-less cars, suppose as a condition of the sale of a driver-less car, state law requires the “car,” not necessarily its purchaser, but maybe a pool consisting of the manufacturer, suppliers, and the purchaser, to obtain insurance sufficient to address any likely issue of liability. To ensure that the car remains insured, the car itself would be the policy-holder, and could not operate without valid insurance. The car might have to have a “kill switch” that would automatically disable the car in the event its insurance lapsed. But the point here is simply that the law could evolve to bestow “personhood” on machines, just as it has done for corporations. See Santa Clara Cnty. v. S. Pac. R. Co., 118 U.S. 394 (1886) (observing that “corporations” may be “persons” for the purposes of the Fourteenth Amendment). My Georgetown colleague, Lawrence B. Solum, first floated this idea more than two decades ago. See Lawrence B. Solum, Legal Personhood for Artificial Intelligences, 70 N.C.L. Rev. 1231 (1992).

28 Gary Marcus, Moral Machines, New Yorker (Nov. 27, 2012), http://www.newyorker.com/online/blogs/newsdesk/2012/11/google-driverless-car-morality.html. As Professor Marcus observes, it may be that truly autonomous military robots or drones will be introduced first. See also Jason Palmer, Call for Debate On Killer Robots, BBC News (Aug. 3, 2009, 7:09 PM), http://news.bbc.co.uk/2/hi/8182003.stm (explaining that the capacity to introduce drones that are able to make firing decisions without direct human oversight will soon be available). But that example does not lend itself to an exploration of the application of product liability rules to autonomous machines, since the law of war will plainly trump the law of products liability. See generally John Markoff, War Machines: Recruiting Robots for Combat, N.Y. Times, Nov. 28, 2010, at A1 [hereinafter Markoff, War Machines]; Joshua Foust, Soon, Drones May Be Able to Make Lethal Decisions on Their Own, Nat'l J. (Oct. 8, 2013), http://www.nationaljournal.com/national-security/soon-drones-may-be-able-to-make-lethal-decisions-on-their-own-20131008.


31 See Funkhouser, supra note 29, at 437-38, 443-44 (discussing the technology that Google cars use to avoid collisions); Smith, supra note 29, at 18-19; Garza, supra note 29, at 587-88.


33 Sven A. Beiker, Legal Aspects of Autonomous Driving, 52 Santa Clara L. Rev. 1145, 1149 (2012) (pointing out that “[d]river error is by far (95%) the most common factor implicated in vehicle accidents”); Markoff, War Machines, supra note 28 (describing the passive role of the “driver” in a self-driving car); see also RAND Report, supra note 30, at 12-16.

34 See, e.g., Beiker, supra note 33, at 1149 (identifying human error as the primary catalyst in the vast majority of car accidents); Funkhouser, supra note 29, at 437-38, 443-44 (discussing Google's attempts--through its self-driving cars--to eliminate human error in driving).

35 Again, this is an assumption, but because the vehicle's control system itself is so highly automated and continually monitors its own performance, it is likely that the vehicle's own data will shed considerable light on the cause of accidents. Indeed, most cars sold today contain “black boxes” that monitor the car's performance and often provide important clues in ascertaining the causes of accidents. See, e.g., Jaclyn Trop, A Black Box for Car Crashes, N.Y. Times, July 21, 2013, at B1, available at http://www.nytimes.com/2013/07/22/business/black-boxes-in-cars-a-question-of-privacy.html. There is a related consideration that is worth noting: Highly autonomous cars will communicate with cars in their vicinity to ensure that a safe distance is maintained between the two vehicles. In that case, it may be that data from other, nearby vehicles may shed light on vehicle failures. See RAND Report, supra note 30, at 66-68, 79-81; Robert B. Kelly & Mark D. Johnson, Defining a Stable, Protected and Secure Spectrum Environment for Autonomous Vehicles, 52 Santa Clara L. Rev. 1271, 1310 (2012) (discussing “vehicle-to-vehicle collision avoidance” communications).

36 See generally Restatement (Third) of Torts: Products Liability §3(a) (1998) (providing for an inference based on concepts of res ipsa loquitur where the plaintiff can show that the product failure “was of a kind that ordinarily occurs as a product defect”); see also In re Toyota Motor Corp. Unintended Acceleration Mktg., Sales Practices, & Prods. Liab. Litig., ___F. Supp. 2d___, 2013 WL 5763178 (C.D. Cal. Oct. 7, 2013); Douglas A. Kysar, The Expectations of Consumers, 103 Colum. L. Rev. 1700, 1721 & nn.89-90 (2003). These sources and others make clear that an inference can take a plaintiff only so far; the plaintiff still has to prove that the failure was of the kind ordinarily seen with design defects. An otherwise inexplicable failure, which is not fairly described as “ordinary,” would likely not qualify under this standard.

37 See generally Larsen v. Gen. Motors Corp., 391 F.2d 495 (8th Cir. 1968) (applying Minnesota law and discussing apportionment of damages in automobile accidents cases); David G. Owen, Products Liability Law §17.4, at 1094-100 (2005).

38 The law generally limits the liability of component part manufacturers absent proof that they participated in the design or manufacture of a finished product. See Davis v. Komatsu Am. Indus. Corp., 42 S.W.3d 34, 43 (Tenn. 2001) (holding that the maker of a robot used in can-making was not liable when the plaintiff's injury was likely caused by unsafe maintenance and other dangerous practices); Kysar, supra note 36, at 1725 & n.109 (and authorities cited therein).

39 This proposal would entail application of “common enterprise” liability. As explained in detail below, I envision this as a doctrine that aims to force a group of companies that work together for a common end--here, to design and manufacture a driver-less car--to bear liability collectively when something goes wrong and injury ensues and when it is impossible to determine, let alone apportion, fault. In other words, as I envision it, common enterprise liability here would be a form of court-compelled insurance. The manufacturers and designers (“the enterprise”) would jointly indemnify individuals injured by driver-less cars when it is impossible to determine fault. In the field of consumer protection, for instance, the Federal Trade Commission often invokes the “common enterprise” doctrine to seek joint and several liability among related companies engaged in fraudulent practices. See, e.g., FTC v. Network Servs. Depot, Inc., 617 F.3d 1127, 1142-43 (9th Cir.
Common enterprise liability should not be confused with its first-cousin, “enterprise liability,” which proposes a collective theory of liability for companies engaged separately in the same hazardous industry, when the identity of the responsible firm cannot be determined. See generally Owen, supra note 37, §11.3, at 752. Compare Hall v. E.I. Du Pont De Nemours & Co., 345 F. Supp. 353 (E.D.N.Y. 1972) (invoking “enterprise” liability theory to hold the highly concentrated blasting cap industry collectively liable for injuries to children), with Sindell v. Abbott Labs, 607 P.2d 924, 935 (Cal. 1980) (rejecting enterprise liability theory). See generally James A. Henderson, Jr., Echoes of Enterprise Liability in Product Design and Marketing Litigation, 87 Cornell L. Rev. 958 (2002); Gregory C. Keating, The Theory of Enterprise Liability and Common Law Strict Liability, 54 Vand. L. Rev. 1285 (2001). See also Kysar, supra note 36, at 1708 & n.28 (collecting sources). Of course, if the number of manufacturers of driver-less vehicles was relatively small, and there were issues of identifying the manufacturer of a vehicle that caused significant harm, enterprise theory of liability might be viable in that situation as well.

According to one press report on the operation of a Google car, “[t] o gain control of the car [the driver] has to do one of three things: hit a red button near his right hand, touch the brake or turn the steering wheel.” Markoff, War Machines, supra note 28. The article reports that during one lengthy test drive, the driver “did so twice, once when a bicyclist ran a red light and again when a car in front stopped and began to back into a parking space. But the car seemed likely to have prevented an accident itself.” Id.

Peter Valdes-Dapena, Thrilled and Bummed by Google's Self-Driving Car, CNN Money (May 18, 2012), http://money.cnn.com/2012/05/17/autos/google-driverless-car; see also RAND Report, supra note 30, at 58-65, 68-69.

This discussion also does not address questions of the crashworthiness of the vehicle, since that question is entirely independent of the identity of the driver.

See generally Owen, supra note 37, Parts I & II. No doubt prompted by Google's development of a wholly autonomous driver-less car, there has been a spate of articles and student notes and comments addressing liability questions. See, e.g., Beiker, supra note 33; Funkhouser, supra note 29; Bryant Walker Smith, supra note 29; Kyle Colonna, Note, Autonomous Cars and Tort Liability, 4 Case W. Res. J. Tech. & Internet 81 (2012); Garza, supra note 29; Julie Goodrich, Comment, Driving Miss Daisy: An Autonomous Chauffeur System, 51 Hous. L. Rev. 265 (2013); Gurney, supra note 29.

On the duty of care, see generally Restatement (Third) of Torts: Physical & Emotional Harm §7(a) (2010) (“[A]n actor ordinarily has a duty to exercise reasonable care when an actor's conduct creates a risk of physical harm.”).

Whether programmers will also be subject to design claims is less clear. Computer code has not generally been considered a “product” but instead is thought of as a “service.” As a result, to the extent that there are cases seeking compensation caused by allegedly defective software, those cases have ordinarily proceeded as breach of warranty cases under the Uniform Commercial Code rather than product liability cases. See, e.g., Motorola Mobility, Inc. v. Myriad France SAS, 850 F. Supp. 2d 878 (N.D. Ill. 2012) (case alleging defective software pleaded as a breach of warranty); In re All Am. Semiconductor, Inc., 490 B.R. 418 (Bankr. S.D. Fla. 2013) (same); Daniel B. Garrie, The Legal Status of Software, 23 J. Marshall J. Computer & Info. L.
The Restatement (Third) of Torts formally abandons the consumer expectations test for design defect claims. See Restatement (Second) of Torts §402A cmt. g (1965); Kysar, supra note 36, at 1712-13.

For instance, in Potter v. Chicago Pneumatic Tool Co. the Connecticut Supreme Court held that the new Restatement permits consideration of consumer expectations, and, in any event, many jurisdictions continue to permit cases to proceed under a consumer expectations theory. See, e.g., Potter v. Chi. Pneumatic Tool Co., 694 A.2d 1319, 1333 (Conn. 1997) (continuing to permit cases to proceed under the consumer expectations test); Kysar, supra note 36, at 1726-29 & nn.111-25; Owen, supra note 54, at 335, 342.

Of course, these scenarios assume that the car was driving at a speed (say twenty miles per hour) within the speed limit and with due attention to the then-current driving conditions (surrounding traffic, road conditions, weather, time of day, and so forth). See generally RAND Report, supra note 30, at 59 (addressing the technological aspects of how a driver-less car would respond to these kinds of scenarios).

I have no intention of joining the debate over the proper formulation of the design defect test; suffice it to say that many eminent tort scholars have waded into this thicket and have emerged arguing for a wide variety of approaches. See, e.g., Richard A. Epstein, Products Liability: The Search for the Middle Ground, 56 N.C. L. Rev. 643, 647-49 (1978) (describing judicial confusion in assessing design defects); W. Page Keeton, Product Liability--Design Hazards and the Meaning of Defect, 10 Cumb. L. Rev. 293, 298 n.23 (1979) (“The search for the universally acceptable definition of defect has been the most elusive one in the products liability field.”); Kysar, supra note 36, at 1709 & n.30; Joseph W. Little, The Place of Consumer Expectations in Product Strict Liability Actions for Defectively Designed Products, 61 Tenn. L. Rev. 1189, 1190 (1994) (“The difficult and politically contentious cases are those that involve allegations of defective design.”); Marshall S. Shapo, In Search of the Law of Products Liability: The ALI Restatement Project, 48 Vand. L. Rev. 631, 638 (1995) (“[A] crucial aspect of products liability law--perhaps the core concept, if any one idea may be described that way--lies in the definition of defect.”); Marshall S. Shapo, Products at the Millennium: Traversing a Transverse Section, 53 S.C. L. Rev. 1031, 1033 (2002) (“However divided analysts of products law may be about definitions, most would agree that the heart of the matter in products liability is the concept of defect.”); David G. Owen, Design Defects, 73 Mo. L. Rev. 291, 292 (2008) (“The quest for understanding design defectiveness perennially vexes courts and accomplished products liability lawyers attempting to unravel design defect problems; delights law clerks, young associates, and law students, furnishing them with an occasion to display their erudition; and provides fertile grist for law professors aspiring for the renown that accompanies discovery of the key to any riddle wrapped in a mystery inside an enigma.”) (footnotes omitted).

See Restatement (Second) of Torts §402A cmt. g (1965); Kysar, supra note 36, at 1712-13.

The Restatement (Third) of Torts formally abandons the consumer expectations test for design defect claims. See Restatement (Third) of Torts: Products Liability §2 cmt. g (1998) (“[C]onsumer expectations do not constitute an independent standard for judging the defectiveness of product designs.”). Many commentators suggest, however, that the risk-utility test adopted by the new Restatement permits consideration of consumer expectations, and, in any event, many jurisdictions continue to permit cases to proceed under a consumer expectations theory. See, e.g., Potter v. Chi. Pneumatic Tool Co., 694 A.2d 1319, 1333 (Conn. 1997) (continuing to permit cases to proceed under the consumer expectations test); Kysar, supra note 36, at 1726-29 & nn.111-25; Owen, supra note 54, at 335, 342.

For instance, in Potter v. Chicago Pneumatic Tool Co. the Connecticut Supreme Court held that although today we continue to adhere to our long-standing rule that a product's defectiveness is to be determined by the expectations of an ordinary consumer, we nevertheless recognize that there may be instances involving complex product designs in which an ordinary consumer may not be able to form expectations of safety.
Accordingly, the Court adopted a new test where complex products are involved to incorporate elements of a risk-utility test. The governing test is as follows:

In determining the reasonable expectations of the ordinary consumer, a number of factors must be considered. The relative cost of the product, the gravity of the potential harm from the claimed defect and the cost and feasibility of eliminating or minimizing the risk may be relevant in a particular case. In other instances the nature of the product or the nature of the claimed defect may make relevant factors otherwise.

Id. (internal quotations omitted); see also Seattle-First Nat'l Bank v. Tabert, 86 Wash. 2d 145, 542 P.2d 774 (1975) (earlier adoption of the same test by the Washington State Supreme Court); Gurney, supra note 29, at 260-62. See generally Owen, supra note 54, at 346-53.

Restatement (Second) of Torts §402A(1) (1965); see also Greenman v. Yuba Power Prods., Inc., 377 P.2d 897, 900 (Cal. 1963) (Traynor, J.) (“A manufacturer is strictly liable in tort when an article he places on the market, knowing that it is to be used without inspection for defects, proves to have a defect that causes injury to a human being.”).

Restatement (Second) of Torts §402A cmt. i; see also Kysar, supra note 36, at 1709 & nn.51-52.


Restatement (Third) of Torts: Products Liability §2(b) (1998).

See Connelly v. Hyundai Motor Co., 351 F.3d 535, 541 (1st Cir. 2003) (noting that, under New Hampshire law, the jury could have reasonably found that the airbag was not defective: “[o]n balance, the benefit to the public of including the overly aggressive airbag system in the Sonata outweighed the danger caused by the airbag system (because the system saved many more lives than it took)’’); Owen, supra note 54, at 330-36.

See Soule v. Gen. Motors Corp., 882 P.2d 298, 308 (Cal. 1994) (noting that the ordinary consumer expectations test is appropriate when the everyday experience of the particular product's users permits the inference that the product did not meet minimum safety expectations).


Among the criticisms leveled at the consumer expectations test is that it is indeterminate, or to put a finer point on it, “so vague as to be lawless.” James A. Henderson, Jr. & Aaron D. Twerski, Achieving Consensus on Defective Product Design, 83 Cornell L. Rev. 867, 882 (1998). The same authors have made the same claim repeatedly. See, e.g., James A. Henderson, Jr. & Aaron D. Twerski, Drug Designs are Different, 111 Yale L.J. 151, 178 (2001).

Beiker, supra note 33, at 1149 (pointing out that driver error is “by far (95%) the most common factor implicated in vehicle accidents’’); Markoff, War Machines, supra note 28; Chunka Mui, Will Auto Insurers Survive Their Collision with Driverless Cars? (Part 6), Forbes (Mar. 28, 2013), http:// www.forbes.com/sites/chunkamui/2013/03/28/will-auto-insurers-survive-their-collision-with-driverless-cars-part-6/.

The law has never treated industry custom as even the de facto standard of care. Eighty years ago, Learned Hand's famous decision in The TJ Hooper, 60 F.2d 737, 740 (2d Cir. 1932), made clear that a negligence standard should not strictly be set by industry custom because “a whole calling may have unduly laged in the adoption of new and available devices.” In most jurisdictions, the defendant is permitted to show that the product's design was “state-of-the-art” (that is, what is technologically feasible, not just industry custom) to rebut claims of defect. See, e.g., Potter v. Chi. Pneumatic Tool Co., 694 A.2d 1319, 1346-47 (Conn. 1997) (collecting cases). Although not conclusive, that evidence is highly probative. Id. Mrs. Arnold might
have to prove that the design, although meeting industry's custom, was not in fact state-of-the-art and further improvements were technologically achievable. See id.

The California Supreme Court in Soule v. General Motors Corp., 882 P.2d 298, 309-10 (Cal. 1994), for example, suggests that using a consumer expectations test to determine liability when “esoteric circumstances” and “complicated design considerations” are present would be improper. Mrs. Arnold's case would likely fall into that category.

See, e.g., Potter, 694 A.2d at 1333.

As the Supreme Court of California drove home in Barker v. Lull Engineering Co., 573 P.2d 443, 451 n.7 (Cal. 1978), relying on the consumer expectations test can undermine incentives for manufacturers to enhance the safety features of their products: “The flaw in the Restatement's analysis, in our view, is that it treats such consumer expectations as a 'ceiling' on a manufacturer's responsibility under strict liability principles, rather than as a 'floor.’” The Barker court formulated the applicable design defect test in this way:

[A] product may be found defective in design, so as to subject a manufacturer to strict liability for resulting injuries, under either of two alternative tests. First, a product may be found defective in design if the plaintiff establishes that the product failed to perform as safely as an ordinary consumer would expect when used in an intended or reasonably foreseeable manner. Second, a product may alternatively be found defective in design if the plaintiff demonstrates that the product's design proximately caused his injury and the defendant fails to establish, in light of the relevant factors, that, on balance, the benefits of the challenged design outweigh the risk of danger inherent in such design.

Id. at 455-56. Barker was explained and reaffirmed in Soule, 882 P.2d at 308 n.4. See also Potter, 694 A.2d at 1333-34.


This balancing test may itself be troublesome because it asks jurors to place values on things that do not come with an attached price tag-- including the “value” of human life, injury and suffering, and the “value” to be gained or sacrificed by design modifications. Making matters worse, jurors not only have to assign values, they then have to weigh them against one another and pick a winner and a loser. For a critique of using economic modeling to make such choices, see generally Frank Ackerman & Lisa Heinzerling, Priceless: On Knowing the Price of Everything and the Value of Nothing (2004).

Although there was much critical commentary on the burdens that the risk-utility test would place on plaintiffs, see, e.g., Kysar, supra note 36, at 1721-22 & n.84 (collecting authorities), the tools for lessening those burdens would not apply here. They include cases (a) where the product defect was one in which the everyday experience of the product's users permits a conclusion that the product's design violated minimum safety assumptions, and is thus defective regardless of expert opinion about the merits of the design, Soule, 882 P.2d at 308 (emphasis in original), or (b) where the product defect was of the kind that “ordinarily occurs” so it is fair to base liability on a res ipsa loquitur theory. See Restatement (Third) of Torts: Products Liability §3(a) (1998) (providing for an inference based on concepts of res ipsa loquitur where the plaintiff can show that the product failure “was of a kind that ordinarily occurs as a result of product defect”); Kysar, supra note 36, at 1721 & nn.89-90.

See supra note 67 and accompanying text on the availability of a “state-of-the-art” defense.

Greenman v. Yuba Power Prods. Inc., 377 P.2d 897 (Cal. 1963); Restatement (Third) of Torts: Products Liability §2(a) (“[A] manufacturing defect [exists] when the product departs from its intended design even though all possible care was exercised in the preparation and marketing of the product...”), In the past, product manufacturers ordinarily were held strictly liable for defectively manufactured products. See generally Restatement (Second) of Torts: Products Liability §402A (1965).

As a general matter, in cases where a product or component is manufactured in a way that fails to comport with the defendant's own intended design specifications, courts routinely find liability. See Richard A. Epstein, Modern Products Liability Law 70 (1980); Kysar, supra note 36, at 1709 & n.29.
There are often questions about whether the manufacturer or service provider that maintained the part is responsible for the defect, but that issue has no bearing here. See supra note 37 and accompanying text. Once driver-less cars enter the market, however, maintenance may become a significant issue. As Anderson's RAND Report observes, many of the key sensors degrade over time, software will need to be updated regularly, and GPS systems will have to be updated constantly to reflect road openings and closings. RAND Report, supra note 30, at 66. Although the car itself will play a key role in continually monitoring the updating and performance of its component parts, other entities will play a role as well in ensuring that these components are maintained and functioning properly, potentially complicating the liability questions. Id.

For example, in Ferguson v. Bombardier Service Corp., 244 F. App'x 944 (11th Cir. 2007), the court rejected a manufacturing defect claim against the manufacturer of an autopilot system in a military cargo plane that suffered a catastrophic crash while the plane was on autopilot. The plaintiffs cited some evidence that the autopilot failed to function properly. But the court found equally credible the defense theory that the plane was improperly loaded, so much so that a strong gust of wind caused the plane to crash—a theory consistent with the information salvaged from the aircraft's flight data recorder. And in Nelson v. American Airlines, Inc., 70 Cal. Rptr. 33 (Cal. Ct. App. 1968), the plaintiff was a passenger on an American Airlines flight who was injured when the plane suddenly descended. One theory was that when the pilots engaged the autopilot, a fault in the autopilot caused the sudden descent. The court applied the doctrine of res ipsa loquitur to find an inference of negligence by American Airlines, but ruled that the inference could be rebutted if American Airlines could show that the autopilot did not cause the accident or that an unpreventable cause triggered the accident. The court said that a defect in the autopilot could have caused the accident, as well as the negligent maintenance of the device. Accordingly, the court reversed the lower court's ruling in favor of American and remanded the action for further proceedings. See also Payne v. ABB Flexible Automation, Inc., No. 96-2248, 1997 WL 311586 (8th Cir. June 9, 1997) (per curiam).

Many of the cases addressing this informational defect are those where airplane manufacturers are alleged to have failed to provide adequate training to pilots in the safe use of their aircraft. See, e.g., Glorvigen v. Cirrus Design Corp., 796 N.W.2d 541 (Minn. Ct. App. 2011) (considering but ultimately rejecting failure-to-train claim); Driver v. Burlington Aviation, Inc., 430 S.E.2d 476 (N.C. Ct. App. 1993); Berkebile v. Brantly Helicopter Corp., 311 A.2d 140, 142 (Pa. Super. Ct. 1973), aff'd, 337 A.2d 893 (Pa. 1975).


See, e.g., Sharon Silke Carty, Toyota's Sudden Acceleration Problem May Have Been Triggered By Tin Whiskers, Huffington Post (Jan. 22, 2012), http://www.huffingtonpost.com/2012/01/21/toyota-sudden-acceleration-tin-whiskers_n_1221076.html (reporting that tiny threads of tin had developed in areas in which they might conduct electricity to the systems that control acceleration, and pointing out that this problem had been “implicated in crippling defects besetting a range of equipment, including communications satellites, pacemakers, missiles and nuclear power plants”); see also In re Toyota Motor Corp., 2013 WL 5763178, at *34-35 (discussing several of the plaintiffs' theories of causation).


Sudden-Acceleration Lawsuits, L.A. Times (Dec. 12, 2013), http://articles.latimes.com/2013/dec/12/business/la-fi-toyota-settlement-20131213. This settlement involves personal injury claims resulting from accidents allegedly caused by sudden acceleration, which have been consolidated in In re Toyota Motor Corp., 2013 WL 5763178. These settlement negotiations are potentially nearing a final resolution. Jaclyn Trop & Ben Prostess, Toyota in Talks on Final Settlements Over Car recalls, N.Y. Times (Feb. 9, 2014), http://www.nytimes.com/2014/02/10/business/toyota-in-talks-on-final-settlements-over-car-recalls.html?hpw&ref=automobiles; see also Estate of Edward W. Knoster v. Ford Motor Co., 200 F. App'x. 106, 114 (3d Cir. 2006) (applying New Jersey law and finding that section three of the Restatement (Third) of Torts preserved the res ipsa loquitur inference so that sometimes when a product fails, “‘common experience’ indicates it would not have done so absent a defect”).

See, e.g., Bloomberg, Toyota Loses First Acceleration Lawsuit, Must Pay $3 Million, Automotive News (Oct. 24, 2013), http://www.autonews.com/article/20131024/OEM11/131029935#axzz2r8ypeVIJ (reporting that an Oklahoma jury awarded $3 million in actual damages and Toyota settled the punitive damage claim for an undisclosed amount in a sudden acceleration case resulting in injury to the driver and the death of a passenger). The court presiding over the multidistrict litigation involving personal injury claims applied this rationale in rejecting Toyota's claim for summary judgment in one of the MDL cases (which may have influenced Toyota's decision to settle all of these claims). The court said:

Toyota's Motion for Summary Judgment is premised on the uncontroverted fact that Plaintiff has been unable to identify a precise software design or manufacturing defect and point to physical or otherwise traceable evidence that the defect actually caused the Camry throttle to open from an idle position to a much wider angle without analog input from the driver via the accelerator pedal. To a lesser extent, it is also premised upon the fact that Plaintiff cannot prove the actual failure of Toyota's fail-safe mechanisms in the Camry on the day of the collision. As explained more fully below, Plaintiff's burden at the summary judgment stage is not so onerous.

Essentially, Toyota asks the Court to conclude that the only reasonable inference that may be drawn from the volumes of evidence proffered by the parties is that Mrs. St. John mistakenly applied the accelerator pedal instead of the brake pedal. The Court cannot so conclude. As Plaintiff points out, and as detailed by the Court more fully below, Mrs. St. John's testimony, together with other evidence, much of it expert evidence, support inferences from which a reasonable jury could conclude that the Camry continued to accelerate and failed to slow or stop despite her application of the brakes.


Of course, one might suppose that if the trailing vehicle were also driver-less, the two cars would be communicating course, position, and speed to one another and the trailing vehicle would thus leave sufficient room between the vehicles to stop safely even in an emergency. See also Marcus, supra note 28 (using a slightly different hypothetical to demonstrate the moral questions raised by driver-less vehicles).

In testing prototypes of war-fighting robots, one study cited as a potential threat modern war robots that--like HAL--could “turn on” their human creators, in part as a mode of self-preservation. See Mick, supra note 23.

Pagallo, supra note 17, at 152-53. See generally Solum, supra note 27.


Although the frequency of automobile crashes is slowly declining in the United States, there were still more than 5.3 million crashes in 2011, resulting in more than 2.2 million injuries and 32,000 fatalities. Alcohol was a factor in 39 percent of the fatal crashes. See RAND Report, supra note 30, at xiv.

Lest there be any doubt, my argument is not based on notions of a “no-fault” liability system, that is, a system that substitutes mandatory insurance and eliminates access to the judicial system. My proposal is a strict liability regime implemented by the courts. Although the idea of “no fault” systems took hold in the 1970s and 1980s, and was expected to drive down insurance costs by limiting the transaction costs related to litigation, it is by now apparent that those systems have not worked as envisioned. See, e.g., James M. Anderson et al., The U.S. Experience with No-Fault Automobile Insurance: A Retrospective, at xiii (2010), available at http://www.rand.org/pubs/monographs/MG860. It is likely, however, that the introduction of driver-
less cars will shift liability from the “driver” to the manufacturer, and that shift may trigger a resurgence of interest in “no fault” insurance regimes.

For that reason, it does not seem fair to saddle the vehicle's owner with the liability. The scenario assumes that the owner's conduct has nothing at all to do with the accident. Nonetheless, if the risk of accidents involving driver-less cars is as low as some experts forecast, then the costs of insuring the vehicle may be so low that it is simply easier for all concerned to make the owner, through mandatory insurance (as is the case in most states already), the responsible party. Moreover, one could easily envision a system where private ownership of vehicles becomes a relic. Instead, companies (or cities or smaller communities) would own fleets of driver-less cars and dispatch them when requested, the way we now use cab services or Uber. This approach could also be used to promote ride-sharing, thereby further reducing the demand for vehicles, fuel costs, and traffic congestion.

See supra notes 37-39 and accompanying text. There is an additional point that might cut in favor of having the manufacturer bear the costs. If one accepts the principles of “Moore's Law”—that computing power doubles every eighteen months—we can expect that the costs of the software components will decrease or remain stable over time while the cost of the other components of the vehicle will continue to rise. If this prediction is borne out, then the manufacturer will increasingly be in the best position to pay, as the software will be an ever decreasing part of the vehicle's overall cost.

FTC v. Tax Club Inc., ___F. Supp. 2d___, 2014 WL 199514, at *5 (S.D.N.Y. Jan. 17, 2014); see also FTC v. Network Servs. Depot, Inc., 617 F.3d 1127, 1142-43 (9th Cir. 2010); SEC v. R.G. Reynolds Enters., Inc., 952 F.2d 1125, 1130 (9th Cir. 1991). To determine whether a common enterprise exists, the court considers factors such as common control; the sharing of office space and officers; whether business is transacted through a maze of interrelated companies; the commingling of corporate funds and failure to maintain separation of companies; unified advertising; and evidence that reveals that no real distinction exists between the corporate defendants. FTC v. Grant Connect, LLC, 827 F. Supp. 2d 1199, 1216 (D. Nev. 2011) (quoting FTC v. Nat'l Urological Grp., Inc., 645 F. Supp. 2d 1167, 1182 (N.D. Ga. 2008)).

There are other liability systems that are similar to the one suggested here. One form of “enterprise liability” is the system underlying the National Childhood Vaccine Injury Act of 1986, which sets up a no-fault system of compensation for children injured by certain vaccines, with funding mainly from the companies that make and sell the vaccines. See 42 U.S.C. §§300aa-1-300aa-34 (2006). The Act imposes a tax on the sale of vaccines to create a fund to pay for injuries attributable to vaccines. See 26 U.S.C. §9510 (2006) (establishing the Vaccine Injury Compensation Trust Fund). If the plaintiff determines that the compensation offered by the program is inadequate, the plaintiff may then file suit against the manufacturer, but most cases are resolved without litigation. See generally Bruesewitz v. Wyeth LLC, ___U.S.____, 131 S. Ct. 1068 (2011).
Abstract

This Article explores the interaction of artificial intelligence (AI) and machine learning with international humanitarian law (IHL) in autonomous weapon systems (AWS). Lawyers and scientists repeatedly express a need for practical and objective substantive guidance on the lawful development of autonomy in weapon systems. This Article proposes five foundational principles to enable development of responsible AWS policy. The findings emerged from a research project conducted by a team of military and civilian professors at the Stockton Center for the Study of International Law at the U.S. Naval War College. The study is informed by experts in computer sciences, government and military, non-governmental organizations, think tanks, and academia.

Advances in AI will likely produce AWS that are different in kind from existing weapon systems and thus require a fresh approach to evaluating IHL compliance. First, this Article describes the technological details pertinent to understanding the distinction between current and future systems. It argues that the technological evaluation of the spectrum of autonomy should focus on the combination of authorities granted to the computer that controls an AWS, while also taking into account the physical capabilities of the system. Second, it argues that a key issue bearing on IHL compliance is whether an AWS has been granted some combination of authorities and capabilities that functionally delegate the decision to kill from human to machine. Third, it posits that predictability must be at the core of an evaluation into whether a particular AWS breaches this delegation threshold and examines how AI handles uncertainty, a critical component of the predictability analysis. Finally, the Article proposes five foundational principles to guide the development of AWS policy.

Table of Contents

I. The Roots of Autonomy and Controversy
   A. Historical Perspective
   B. Framing the Discussion
II. Autonomy, Artificial Intelligence, and Machine Learning
   A. An IHL Analysis of AWS Should Focus on the Computer System
   B. Describing Autonomy from a Technical Perspective
   1. The Spectrum of Autonomy
Autonomous weapon systems (AWS) are the most militarily significant yet legally elusive challenge to international humanitarian law (IHL) since the proliferation of cyber operations. The modern debate over AWS ignited following the release of Losing Humanity, a report co-authored by Harvard's International Human Rights Clinic and Human Rights Watch. Since then, academics, government officials, non-government organizations (NGOs), and military leaders alike have struggled to address the myriad legal concerns potentially raised by AWS.

In the years following Losing Humanity, it became apparent that the dilemmas presented by AWS would never be solved by any one professional field operating in isolation. Lawyers were hamstrung by a dearth of technical expertise. Scientists were hampered by a lack of legal acumen. Non-military personnel were confounded by their unfamiliarity with the likely battlefield application of AWS. And everybody was forced to sift through a significant amount of misinformation surrounding AWS.

But after years of debate, we still continue to ponder: is there nothing to be concerned about, as some in government and industry would have us believe? Or is humanity's war with machines imminent unless we take immediate action to ban AWS completely, as a few NGOs argue? And, more to the point, why is applying IHL to AWS so difficult? After countless academic discussions, law review articles, conferences, meetings of experts, and consultations between state representatives, the conversation about AWS has matured little. There remains a dearth of practical guidance on how states should regulate AWS development. The root cause underlying this lack of progress is simpler than one might suspect: IHL is a sub-optimal tool for remedying our inability to predict the future.
IHL governs the conduct of parties to a conflict. Its principles and rules generally serve to balance the practical realities of armed conflict with the desire to protect civilians from harm and combatants against unnecessary suffering. Under IHL, attacks must serve a valid military purpose. A target may not be attacked unless it qualifies as a military objective and the commander must weigh the military advantage to be gained in attacking a target against expected collateral damage to ensure the collateral damage is not excessive in the circumstances. The principles of IHL have developed as a function of treaty and customary international law distilled from the lessons of countless armed conflicts. It is tempting to think that we might use IHL to completely forestall some yet-undefined harmful technology. States toil tirelessly, however, to resolve legal problems that already exist; when it comes to solving problems that are not yet realized, the task is nearly insurmountable. But why must we predict the future--could we not simply inventory all potential AWS and then make some general statements about how IHL might apply? Probably not. An inventory approach would have to determine, as a preliminary matter, precisely what constitutes an AWS. Are we referring to existing weapon systems that exhibit a degree of autonomy, such as the Phalanx close-in weapon system (CIWS) and Counter Rocket, Artillery, and Mortar (C-RAM) systems? Or to likely near-future systems, such as unmanned sub-hunting ships or swarming mini-drones? Alternatively, are we referring to some fanciful, unrealized future concept such as killer robots? The answer is that we simply do not know what to inventory because in many instances we would be trying to predict the development of systems that do not exist, are unlikely to exist in the foreseeable future, and might never exist. This is not how the law usually develops.

Law, and IHL in particular, is typically reactive. Generally speaking, there are good reasons for this. Humans cannot predict the future. Attempting to devise complex legal schemes that will effectively anticipate actions and technologies that do not exist is vexing at best and likely counterproductive. As such, the law is more adept at fixing things that are broken after careful discussion and debate.

In the context of IHL practice, however, this means that a lot of people perish before a particular problem is addressed. As such, with the development of any weapon there is an unavoidable tension. On one hand, we may seek to prevent needless death and suffering by generally restricting the types of weapons produced and also by specifically tailoring the lethal effects of new weapons to fit a certain purpose. On the other hand, in the interest of protecting our national security we may seek to create new weapon systems that can eliminate enemy threats, which historically has meant that they were overwhelmingly destructive. From the second perspective, technological advantage is maintained by legal review of new systems on a case-by-case basis, placing concomitant limitations on use rather than broader restrictions on procurement. There is no simple resolution to this dilemma--most parties involved in the development of IHL strive to strike a reasonable balance between these competing interests.

The potential impact of AWS technology on this balance is nebulous, but we must nevertheless glean meaningful practical guidance. We should not attempt, however, to solve intractable problems by debating capabilities that may not exist even 100 years from now; those matters are simply too speculative to provide a foundation for meaningful debate. We should instead evaluate from a technical perspective the particularized challenges presented by autonomy in the reasonably foreseeable future. This Article seeks to craft broad but useful and substantively meaningful principles that further an understanding of IHL's application to AWS. In particular, the Article will dissect AI and its interplay with IHL.
This Article proceeds initially with a brief historical perspective on issues surrounding autonomy and then pinpoints the details pertinent to understanding how AWS technology intersects with IHL. After thus framing the discussion, it explains why an IHL analysis of AWS must focus primarily on the computer system. The Article next explains why future autonomy should be related logically to the human decision-making cycle, then explores how AI works and, in particular, how machines learn. It argues that in the context of AWS equipped with AI, an IHL evaluation must focus on predictability, and then examines how AI handles a primary challenge to predictability: uncertainty. After a brief review of the relevant IHL concepts, the Article proposes five principles to guide the development of AWS technology.

I. The Roots of Autonomy and Controversy

A. Historical Perspective

Automation is not new. Since the dawn of civilization, human beings have sought tools to mechanically assist them in completing a myriad of tasks. In every endeavor across the spectrum of the human experience, mankind has developed implements designed to alleviate burdens previously borne by people. From agriculture and industry to national defense and warfare, thousands of years of technological development have yielded devices that preceding generations could not have imagined. Mechanization led to automation that helped us to accomplish our goals not only with greater ease, but also more quickly and efficiently than was previously possible. Unsurprisingly, humans have invariably adapted new technology to military applications when it might provide a warfighting advantage.

The pace of automation has increased exponentially. Machines have transitioned from simple automatic implements fashioned to assist us to autonomous computerized mechanisms sometimes capable of replacing us. Konrad Zuse invented the first operational programmable computer in 1941. Today--within the span of one human lifetime--automobiles can drive themselves. Bursts of technological advancement have induced speculation about whether machines might eventually outperform their creators not only in routine physical tasks but also in behaviors previously viewed as uniquely human. The core concern about autonomy is that we might eventually design machines so advanced that they slip out of the grasp of human decision-making and control.

B. Framing the Discussion

Decisions regarding the use of force in any context are often quite controversial. But in order to have an informed debate about these matters in the context of AWS it is first necessary to examine carefully, from a technological perspective, what it means to “decide.” In other words, what does it mean to say that a human decided that a machine would accomplish a given task? Conversely, when is control so attenuated that it could no longer reasonably be said that a human decided a machine would accomplish that task or the manner in which the machine should complete it?

One decision in particular--the decision to kill--lies at the heart of concerns over AWS. The decision to kill inherently invokes analysis under IHL as to the lawfulness of a use of force. The burden of conducting this evaluation logically and necessarily must be borne by a human. But the link between a human's decision to kill and the lethal kinetic action of a weapon continues to steadily degrade as a function of proximate cause, and in particular as a temporal matter. Thus, the appraisal of whether a human decided to kill is not a digression into philosophical inquiry; in the AWS context it is instead a technological assessment. We must determine whether AWS technology could unlawfully
dilute this causal link such that we could no longer say that a human functionally decided to kill.\textsuperscript{45} Note, however, that this is not meant to imply that a human must provide an AWS input that is temporally proximate to lethal kinetic action.

As a corollary, we must scrutinize from a technological perspective the concept of what it means to “control” a machine. Before we could hope to answer the question of what amount or kind of human control over AWS is legally sufficient,\textsuperscript{46} we must first possess a firm understanding of how humans control machines via programming. Indeed, if we misapprehend the manner in which machines are controlled, for example by layering legal significance onto the proximity of human interaction with the machine at the time of lethal kinetic action, then we risk misleading ourselves. Only after the technical underpinnings of machine decision-making\textsuperscript{47} and control have been explored can we hope to parse out aspects that might prompt a legal objection. Thus, we must describe from a technological standpoint how machines “decide” and how humans control those “decisions.”

\textbf{II. Autonomy, Artificial Intelligence, and Machine Learning}

\textbf{A. An IHL Analysis of AWS Should Focus on the Computer System}

As a point of departure, consider a conventional military unit: an artillery battery. In evaluating the ability of the battery to comply with IHL, one might start by investigating the performance characteristics of the cannons. Testing the accuracy of cannons is relatively easy.\textsuperscript{48} But more to the point, the cannons do not decide where to aim themselves. A modern howitzer receives firing data from a computer system (a fire-control computer)\textsuperscript{49} which has a straightforward task: given a set of geographical coordinates, calculate the proper elevation and deflection settings for the cannon while taking into account meteorological conditions and other measurable factors that will affect the trajectory of a round.\textsuperscript{50} If the firing data caused the cannon to hit the target, it was correct.\textsuperscript{51} If the round was off target, the data was incorrect. This is simple computer automation. The battery commander is responsible for ensuring that his or her weapons are employed lawfully. It would be inconceivable for a prosecutor to say, for example, that an IHL violation was the fault of the fire-control computer. The computer system is a very advanced calculator. But for AWS, the computer system plays a central role in the IHL analysis.

In the context of AWS, the primary focus of technological analysis necessarily shifts to the integrated system rather than the weapon component alone, for two reasons. First, computer-managed systems enable weapons to be removed from the continuous physical control of a human; a weapon that is not incorporated into such a system is therefore unremarkable from an IHL-compliance perspective because it ceases to be an AWS.\textsuperscript{52} Second, it is a safe assumption that autonomy will continue to increase in modern weapons.\textsuperscript{53} As such, the computer systems linked to weapons will play an increasingly significant role in how the weapon is employed. With respect to the system, we focus primarily on the computer that effects control of the machine. This is because, in large part, the level of autonomy a system enjoys is determined by the computer that effects control of it.

To be sure, the physical capabilities of the mechanical platform on which the computer is installed, as well as the characteristics of any accompanying weapon may play a significant role in describing the autonomy of the complete system. But if the computer effecting control of the machine sets limiting parameters on the system, the overall capabilities of the system may be moot. The converse is not true. By way of simple example, an unmanned aircraft might have the mechanical ability to carry a large unguided bomb. But if the sophisticated targeting computer onboard the aircraft only allows it to vector towards unpopulated areas in order to attack positively identified enemy tanks during an international armed conflict, our concerns over civilian casualties may be reduced. On the other hand, a similar platform with a
different computer system that is granted unbridled discretion in attacking targets could raise significant legal concerns, almost without regard to the type of weapons on board. Indeed, it is difficult to imagine any weapon system that would not appear objectionable if given “a mind of its own” in the most human sense.

However, machines do not have minds of their own—they have computers programmed by humans. Generally, computers do what they are told to do. That said, computers of today are infinitely more complicated than they were even a few decades ago. Concepts such as machine learning and AI are quickly becoming the focus of the discussion on autonomy, and factors like these greatly complicate the control analysis. Is it possible that advanced machines could eventually “decide” to do something other than what a human “told” them to do, or do something that we thought them incapable of doing?

We must therefore delve deeper into the technological manner in which autonomy functions in machines and the role that machine learning and AI are likely to play in its future. Only with this understanding can we fashion a legal framework that adequately addresses concerns about control over machine decision-making.

*392 B. Describing Autonomy from a Technical Perspective

1. The Spectrum of Autonomy

The notion of autonomy in the context of machines is so broad as to defy simple definition. It is accordingly helpful to think of autonomy as a spectrum or series of spectrums rather than a singular concept. And as described in the previous section, autonomy with respect to AWS must concentrate primarily on the computer system while also taking into account the physical capabilities of the system. As such, a description of the AWS spectrum should focus on the combination of authorities and capabilities that designers grant to computer systems running AWS. There is no flawless analytical construct to accomplish this task, but a helpful method for describing potential combinations is the Boyd Cycle, or “OODA Loop.”

The OODA Loop is a simple way to evaluate human decision-making based on a continuous process: Observe, Orient, Decide, Act. As part of the cycle:

- a person first observes the world around her, gathering data about her environment through the array of human senses. Second, she orients herself, or interprets the information she has gathered. Third, she weighs the potential courses of action based on the knowledge she has accumulated and decides how to act.
- Fourth and finally, she acts, or executes the decision she has made.

Through this model, we can generally describe the types of authorities that might be granted to a machine that could supplant the human’s role in fulfilling the requirements of the various points in a decision-making cycle. Importantly, in the context of controlling AWS, this Article does not refer to the moment where a human might be inserted into the cycle, but instead to those fragments of the loop that have been delegated to computers. Granting certain portions of the OODA loop to machines may ultimately create issues with IHL compliance. This key distinction is illustrated in the following figure:

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE
In this vision of the OODA loop, the puzzle pieces of tasks delegated to a computer in the future might consist of authority (e.g., in the precise programming or learning capacity of the computer) and/or physical capabilities (e.g., the ability of the host platform to loiter for long duration). Thus, the key issue bearing on IHL compliance is not whether the machine, for example, selects and engages targets (i.e., decides and acts) without human intervention. Rather, the critical issue is whether designers have granted the machine some combination of tasks that functionally delegates the decision to kill from human to machine.

2. The “Observe” Phase

Machines have historically performed the task of observing military operations and they continue to advance in their capability to do so. The first airplanes were used as platforms for humans to reconnoiter First World War battlefields. Now, unmanned vehicles carry sensor suites that allow humans to observe the enemy from afar. Ground-based sensor platforms guard demilitarized zones from intrusion and ship-mounted sensors can detect and destroy threats to warships. Facial recognition technology and other sensors may soon advance to the point where machines are able to positively identify a person without additional input by a human at the time of observation. Indeed, the range of sensors deployed in today's modern militaries is vast.

Nevertheless, machine observation and identification of objects and persons on the battlefield have generated little debate or concern from a legal perspective. Indeed, it would be beyond cavil to suggest that IHL restricts machine sensors that simply detect objects. Other fragments in the OODA loop, especially in combination with portions of the Observe phase, may however become troublesome.

3. The “Orient” Phase

In this phase, information gained during the Observe phase is analyzed in order to better understand the operating environment. In the context of human decision-making, a commander will consider the totality of the information at his or her disposal. Current intelligence estimates, sensor collection and battlefield reports are reviewed, the tactical and strategic implications weighed, as are countless other military and non-military considerations. The experiences of the commander play a key role. In the end, the human decision-maker may simply trust his or her “gut” feeling.

Machines do not have guts. They analyze data based on their programming. The process by which a computer-controlled machine completes this task depends on the attributes granted to it by a human. At the most basic level, machines analyze a rather narrow category of information against set decision-making processes. This is essentially a flow chart, or an “if this, then that” method, of evaluating data. Broadly speaking, this process works well for simple machines completing basic tasks.

So far as we can reasonably tell at this time, machines in the foreseeable future will not have guts, hunches, or any of the other qualities unique to humans. They may, however, complete tasks in ways that are far more difficult for humans to predict. For example, the technology embedded even within many currently existing computers contains so many lines of code that testing out all defects can be impractical or even impossible. Machine learning further complicates the task of predicting machine behavior. While the equations employed in programming a machine may themselves be easy to understand, we may not be able to determine ahead of time the results that they will produce. These are the computing methods that will likely be employed in future AWS.
Machine learning and AI therefore pose difficult questions about how machines will perform in the future. Part of the dilemma in this regard exists because humans have proven rather inept at predicting technological advances, even in the short term. With current technologies we can often predict at least what processes a computer will use to analyze data or reverse engineer the process that brought a computer to a certain course of action based on its programming. This capability, however, is rapidly eroding and may not generally be the case in the very near future.

4. The “Decide” Phase

The most provocative and controversial aspect of task delegation to AWS is the Decide phase. This is because humans are generally uncomfortable with machines completing the “final” deliberative step in a sequence of events that will ultimately result in the death of a human. But it is a common error to posit that if a computer takes the last step in a process, the machine “decided” to kill. It does not necessarily follow that a machine “decided” to kill simply because a machine selected and engaged a target based on certain narrow and predetermined parameters. Machines generally do as they are programmed to do, and humans make decisions to program machines a certain way. The programming could be based on immensely complex learning algorithms operating in stochastic environments or it could be simple AI executing trivial tasks in a controlled situation. In the end, actions by the machine are the result of a human delegating a task to the machine for performance according to certain pre-established performance measures. Computers simply do not act on their own volition.

Nevertheless, this is an area that requires additional scrutiny because we must consider whether the link between the programming and lethal kinetic action might become so diluted that we cannot reasonably say a human decided to kill. Further, we must consider the possibility that a machine might be granted such advanced technology that it unexpectedly exceeds anticipated operating parameters or otherwise behaves incongruently with the discretion that humans programmed it to demonstrate. These crucial questions of predictability will be discussed in greater detail below.

5. The “Act” Phase

The final phase in the cycle relates to the physical authority in time and space that designers have granted to the machine. Although this aspect is of secondary significance to the decision-making authority the machine has been granted, physical authorities may prove significant depending on the other aspects of the system. For example, to what kind of weapons is the system provided access? How long is the system able to loiter in the operating environment? How far is it able to travel? Can it detect the presence of nonmilitary objects and/or noncombatants? These and other questions may prove highly relevant to the evaluation of AWS. Then again, lack of discretion on the part of the machine to take advantage of these capabilities may obviate their significance.

In sum, our technological evaluation of the spectrum of autonomy should focus on the combination of authorities granted to the computer system that controls the machine while also taking into account the physical capabilities of the AWS. And the critical issue bearing on IHL compliance from a technological perspective is whether the AWS has been granted some combination of capabilities that functionally delegates the decision to kill from human to machine. This, however, is only the initial step in delineating the kind of autonomy that could prove legally objectionable under IHL. The next step is to more fully explain how autonomy functions from a technical standpoint so that we can fully understand machine decision-making.
C. An IHL Analysis of the Computer System Should Focus on AI

In order to discern the aspects unique to AWS that may interfere with IHL compliance, we must first identify the particular qualities of AI most relevant to the inquiry. In order to outpace sophisticated future adversaries, computers that run AWS may be equipped with the capacity to “learn” even after they are employed by a battlefield commander. As a result, at least certain aspects of future learning-equipped AWS will adapt in ways that we may be unable to predict. As such, historic examples of weapon systems that incorporated AI, which did not employ learning AI, are of little use in this context as future AI will be different in kind from past AI.

1. Surveying Current Technologies is Unhelpful

This Article does not attempt to conduct a comprehensive survey of existing technologies in an effort to craft generalized legal principles for AWS. It is tempting to select various weapons systems currently in existence and then try to rationalize them under some overarching rubric for AWS. One might reasonably question why we cannot simply construct a historically focused *398 framework for systems that might exist in the future. This approach, while sensible at first glance, ultimately would be a fruitless endeavor.

As discussed in Part II.B.1, the breadth of legacy weapon systems--those already deemed lawful and fielded--that could be considered autonomous is vast. The question, for example, of whether or not a conventional land mine is autonomous is intriguing from a purely academic perspective, but the discussion has almost no practical value for those crafting regulations for future AWS development. Other systems that exhibit some aspects of autonomy have also been reviewed and have enjoyed lawful status under IHL for quite some time. As such, even those organizations most vehemently opposed to the development of AWS do not argue that legacy systems exhibiting elements of autonomy would be included in a ban. Thus, an attempt to neatly categorize scores of legacy systems within comprehensive legal principles on AWS would do little to advance the conversation.

More importantly, legacy technology will not significantly inform our evaluation of whether future AWS are lawful under IHL because future systems will be different in kind, not simply different in degree of autonomy. Thus, past *399 analyses regarding the ability of legacy systems to comply with IHL will not be particularly useful in predicting whether future systems are lawful. It goes without saying that legacy systems are generally used as precedent for evaluating new systems insofar as they are reasonably similar. Weapons systems markedly different in kind from legacy systems, however, require novel approaches to applying the law. Weapon systems that are granted advanced AI and learning capability are so different from legacy systems that they require a fresh approach to applying IHL. The focus must thus necessarily be on the systems that incorporate AI and machine learning in new and profoundly different ways.

2. Advanced AI will create autonomy that is different in kind

If left without practical guidance on the lawful development of AWS, future systems might be designed with AI that is so advanced that designers could not predict to a reasonable certainty how it would perform in an operational environment. To say that there is disagreement in the scientific community regarding what paths AI might take in the future would be an understatement. This Article does not enter the fray in this regard. Instead, we begin by establishing a baseline description for what direction--based on extensive research and interviews with leading experts in the field--AI will not take.
Contrary to the news headlines or Hollywood productions that one might encounter when researching AI, \textsuperscript{83} the singularity is not near, \textsuperscript{84} nor is “Skynet.” \textsuperscript{85} *400 Even if such a technological tipping point might theoretically be reached, \textsuperscript{86} IHL typically does not attempt to regulate technology that \textit{might} exist in 30 years, 100 years, or may never exist. \textsuperscript{87} Yet, the question remains whether there is anything that we should be concerned about when incorporating AI into AWS? The answer is yes.

Problems may arise with systems that are quite advanced but not smart enough to do what we demand of them. Specifically, AWS that either fail to meet performance standards under IHL or, more pertinent to the present discussion, whose performance cannot be adequately predicted due to their AI raise very real concerns. \textsuperscript{88} With regard to the former, this is a relatively simple weapons testing question with which countries executing weapons reviews are quite familiar. \textsuperscript{89} But with respect to AI in particular, a significant possibility exists that computers \textsuperscript{401} could advance to the point where we are unable to reasonably predict whether the weapon system will comply with IHL. In other words, a system could be so advanced that we would be unable to conduct testing sufficient to assure us to a reasonable degree of certainty about how it would perform in certain operational circumstances. \textsuperscript{90} For this reason, we must understand how AI works and in what ways predictability and uncertainty manifest themselves in AI systems.

3. The Focus of AI in AWS: Performing to Rational Objective Standards

AI in AWS should be evaluated based upon how well it performs to rational and objective standards. Not all visions of AI follow this construct. Some models of AI are concerned with thought processes and reasoning by computers whereas others focus on behavior and performance. \textsuperscript{91} How we define AI therefore hinges on whether we seek to have machines think like humans, act like humans, think rationally, or act rationally. \textsuperscript{92} In approaching AWS from a legal standpoint, this Article’s primary concern is the ability of a given system to comply with IHL. The focus of any evaluation under IHL is the lethal and destructive effects caused by the AWS. \textsuperscript{93} Thus, the goal is to have the AWS act rationally by producing these effects in accordance with objective standards.

From early in the debate on AWS, most experts have agreed that aspiring to a human performance standard probably set the bar too low. \textsuperscript{94} Indeed, humans sometimes demonstrate a sub-optimal track record in adhering to the standards of IHL. \textsuperscript{95} And most agree that autonomous systems in non-military applications should also be required to outperform humans in rule compliance. \textsuperscript{96} So we proceed from what is likely the uncontroversial proposition that AI in the AWS context should be designed and judged based upon how well the system is able to perform to rational and objective standards established by humans and informed \textsuperscript{402} by IHL. The next question in the inquiry, then, is: What aspects of AI could help AWS achieve (or fall short) of these standards?

4. How AI Works in Practice

The science of AI that enables the design of systems that can behave rationally based on objective performance standards focuses on the programming of an intelligent “agent.” A few terms of art must be explained at this juncture. “An agent is something that perceives and acts in an environment.” \textsuperscript{97} Think of this as the complete system--the physical platform mated with its computer hardware and software. The “agent function” describes what the agent will do in response to given inputs. \textsuperscript{98} The function could range from simple “if/then” logic to nondeterministic functions based on complex algorithmic processes. “Agent programs” implement the agent function. \textsuperscript{99} This is the software programmed into the
system that is designed to achieve the desired (rational and objective) outcome. Programs are tailored to respond to the unique environment in which the agent operates. Agents can also be granted, via programming, the ability to “learn” through their perceptions by adjusting their behavior in order to better achieve assigned goals.

Consider the simple case of the Roomba. The Roomba is a small vacuum that is designed to clean floors without direct physical manipulation by a human. This agent uses a suite of sensors that helps the agent function determine how to navigate and locate objects so as to most efficiently vacuum, then “remembers” where those obstacles were and avoids them in the future. The agent “knows” when its battery is low and the agent program guides the Roomba back to a charging station. The Roomba also can detect whether it is cleaning carpets or hardwood and adjust power output accordingly. With these capabilities one could argue that the Roomba is a fully autonomous and artificially intelligent learning robot.

That said, it is important to avoid the temptation of ascribing broad and unrealistic capabilities to AI simply because a system is quite advanced in certain narrow respects. Careful attention must be paid to what the AI cannot do. Despite some impressive features, the Roomba is constrained in what it can achieve due to its limited agent platform, sensors, and agent program. In other words, if you expect it to do anything other than vacuum the floor, you will be disappointed. More advanced but nevertheless relatively narrow AI-enabled machines are more illustrative and logically relatable to AWS. For example, computers have beaten the best human opponents in complex games such as chess, Go, and the television game-show Jeopardy. While these achievements are impressive, it is important to note the narrow operating environment in which these agents functioned. Simply because an agent is adept at playing a very complex game it does not necessarily follow that its AI is sophisticated enough to conduct high-level reasoning comparable to humans.

As a result, some who criticize the introduction of AI into weapon systems argue that such narrowness in current AI means that it could never be sophisticated enough to handle the complexities of the modern battlefield. Others argue that future AWS will need to cope with decision-making speeds that are beyond human capacity because future wars will be fought at “machine speed.” Many agree, however, that regardless of the resolution of these issues, the sophistication of AI will need to advance in order for AWS to account for or mitigate these realities. The implication of this debate is that in the future, AWS will likely be called upon to “learn” in order to handle complex and changing environments.

5. How Machines Learn

Agents learn by being provided data sets from which an onboard algorithm can be programmed to attain rational goals. The agent may then be placed in an unknown environment in which it will draw upon its learning data sets and endeavor to achieve optimal results. A learning agent will continue to refine its behavior based on the results that it achieves in real-world operating environments as compared to established goals.

Learning agents are generally comprised of a performance element and a learning element. The performance element senses the environment of the agent and determines a course of action, while the learning element employs feedback from the system “on how the agent is doing and determines how the performance element should be modified to do better in the future.” Machine learning in artificially intelligent “agents can be summarized as a process of modification of
each component of the agent to bring the components into closer agreement with the available feedback information, thereby improving the overall performance of the agent.”

By way of basic example, consider a hypothetical stealth drone equipped with learning AI that helps the system avoid detection by enemy radar. The performance element of the drone determines the heading, speed, and altitude at which it will fly. Suppose the drone begins its journey into enemy territory by flying at a high altitude. Enemy radar then detects the craft despite its stealthy design. A “critic” will inform the learning element that the enemy radar is targeting the drone, and the learning element may then inform the performance element that flying at a high altitude is not optimal. The performance element flies the drone to a lower altitude, and it does not again return to higher altitudes. The machine has learned.

In order to fully grasp how intelligent systems learn about the world into which they are placed, it is also necessary to appreciate how agents are programmed to view their environment. The most basic agents are only able to observe their environment in binary terms, in what is referred to as an “atomic” representation of the world. For example, to return to the Roomba hypothetical, assume that the agent is only able to ascertain two states of the world, where the floors are either “clean” or “dirty.” The agent does not consider other factors, such as whether the homeowner might be annoyed by the presence of the Roomba. If the floor is dirty, the Roomba will endeavor to make it clean. There are no other factors for the Roomba to weigh in the calculus.

But suppose we change the hypothetical such that the Roomba is able to measure the tone of the homeowner's voice. Also, what if it knew that houseguests were visiting who might not appreciate its presence? There are a litany of other factors that might bear on the Roomba's ultimate “to clean or not to clean” decision. If the agent were able to sense and weigh these factors, it would be considering either a “factored” or “structured” representation of its environment. When an agent uses a factored representation of its world, it can consider a range of variables and attributes that have discernible values. In a structured representation, the agent is able not only to consider these variables but also to weigh the inter-relationship between them.

A machine's ability to learn using factored and structured representations of the world may be crucial to the application of AI to AWS for two reasons. First, AWS that are fielded will often need to consider multiple variables regarding their operating environment. In certain circumstances, the AWS will need to be able not only to weigh numerous variables that it senses, but it might also need to consider the way in which those variables relate to others. These considerations will continue to impact a wide swath of actions by AWS, such as navigation, object recognition, and fire-control solutions. Second, and more importantly with regard to future AWS, an agent's ability to effectively represent its environment and account for complex interrelated variables while learning may be pivotal in handling uncertainty, as well as ensuring predictability.

It is important, however, not to overstate this point. While future AWS will likely need to be programmed to consider factored or structured representations of their operating environment, it does not follow that they will always need to consider all of the factors a human would consider or be able to make fine-grained judgments about how each relevant variable in the environment relates to the others. This was the fatal assumption of Losing Humanity, as it envisioned a world where AWS would necessarily be called upon to directly substitute for human soldiers in the most complex battlespaces. The simple resolution to this issue has three aspects. First, for most scientists experienced in AI, it is nearly impossible to rationally envision that machines will possess the technological capability to make such complex and subjective decisions, even if equipped with highly sophisticated AI. Second, machines are not compelled to act under
the same time constraints as humans because they have no self-preservation instinct. Third, if a military commander is not reasonably certain that a weapon system will comply with IHL as employed, he is under a positive obligation not to use it. This is true for any weapon or system, past, present, and future.

That said, learning in intelligent systems that are provided lethal capability nevertheless raises significant issues that must be carefully scrutinized. This is because the further we move from deterministic (“if/then”) systems and towards complex learning systems, the less reasonably we may be able to predict how the system will arrive at given solutions. This may or may not present issues with IHL compliance; it depends on which puzzle-shaped pieces of the OODA loop the system has been granted. Each time that a new AWS is designed, we must evaluate whether the learning capacity that it is granted will prevent the AWS from being employed in conformance with IHL.

One might reasonably ask, then, why we would grant an AWS any learning capacity in the first place. Why not simply design the system to account for all of the difficulties that it might encounter in its operating environment? Depending on the environment and the mission of the system, this goal may be impossible. The reasons for this are threefold:

First, the designers cannot anticipate all possible situations that the agent might find itself in. For example, a robot designed to navigate mazes must learn the layout of each new maze it encounters. Second, the designers cannot anticipate all changes over time; a program designed to predict tomorrow’s stock market prices must learn to adapt when conditions change from boom to bust. Third, sometimes human programmers have no idea how to program a solution themselves. For example, most people are good at recognizing the faces of family members, but even the best programmers are unable to program a computer to accomplish that task, except by using learning algorithms.

In essence, then, if we develop AWS that will be able to fight at machine speed in future conflicts by leveraging learning technologies in AI, then we must accept that some aspects of these systems will adapt in ways that we may be unable to predict. As such, we must delineate the implications of this increased uncertainty.

D. Within the Learning-enabled AI, the Analytical Focus is on Predictability

Given the concerns over AI and machine learning, how do we prevent ourselves from functionally delegating those decisions (e.g., the decision to kill) that we may not delegate? Predictability is the key. The analysis here focuses on aspects of a system that might, in combination, affect our ability to reasonably predict its compliance with IHL. It cannot be overstated, however, that not all aspects of the system must be predictable. There is of course great potential military advantage to be gained by providing advanced machine learning, for example, to aspects of a machine that either do not bear on IHL compliance or do not combine with other autonomous features to functionally delegate the decision to kill.

Like most IHL requirements, our ability to predict the actions of the machine must be based on a reasonableness standard. The test of reasonableness emanates from the recognition under IHL that attaining even near-certainty in armed conflict is usually an insurmountable goal. From a practical standpoint, a lower standard would encourage noncompliance with IHL by inviting humans simply to blame erratic computers for violations. A higher standard would likely be unattainable based on the complexity of computer programming magnified by the “fog” of the modern battlefield. For a variety of policy reasons, states may want to set higher bars for their own introduction of certain
AWS into operation. As a matter of law, however, the reasonableness standard is a well-established benchmark of performance that has balanced the competing interests of IHL for quite some time.

Predictability cannot diminish past the point where we can reasonably say a human was in control of compliance with IHL. Importantly, this is not the same standard as physical human control over the actions of the machine itself at the time of lethal kinetic action. It also does not mean that a human made a call on IHL compliance that was temporally proximate to a lethal attack. Rather, it means that we can reasonably predict what action the system will take and that we are reasonably certain that the system will comply with IHL. If we can reasonably predict compliance, then we maintain control no matter the level or type of our interaction with the machine at the instant of lethal action. But if we cannot reasonably predict whether the machine will comply with IHL, it may be unlawfully autonomous.

III. Machine Learning and Predictability

A. Uncertainty in AI and the Pitfalls of Unpredictability

AWS must be designed to account for at least some of the uncertainty deriving from complex operating environments. In order to evaluate the ability of a system to meet this requirement we must understand from a technical perspective how uncertainty is handled. Generally speaking, AI accounts for uncertainty by weighing the probability of certain outcomes against the desirability of such outcomes. The ability of the system to meet performance goals in the face of uncertainty is a key variable in these calculations.

When we ask whether a particular system equipped with AI is lawful per se and whether it can comply with IHL as employed on the battlefield, we must necessarily inquire into whether it can meet technical performance standards. We know that these performance standards must be based on rationally described goals or endstates. In controlled laboratory environments, rational performance standards might be easy to describe. The environment in which the AI is called upon to operate could be quite simplistic. This makes the task of designing and testing that sort of AI easier. But what about on a battlefield?

Even the most straightforward combat scenarios are often immensely more complex than a laboratory environment. The challenges computer programmers face in designing AI that can handle uncertain environments are significant. A program that is designed to handle all possible eventualities presented even in environments of relatively limited uncertainty could require impossibly large data sets. This has led some groups to conclude that AI could never function effectively in combat while still adhering to IHL standards. The weakness of such critiques, however, is that they assume too much about how AI will function on the future battlefield. In fact, most people--regardless of their ideological or institutional biases--assume far too much about what roles machines might fulfill and how they might go about fulfilling them. The root cause of these assumptions is our collective concern over how future AI will handle uncertainty because we worry that machines will perform unpredictably.

But science fiction aside, even the most advanced computers and cutting edge AI should perform to some level of predictability. If we carefully instruct computers how to account for and respond to uncertainty, we should in theory be able to predict IHL compliance (or lack thereof) to a reasonable certainty. Again, this will depend heavily on the pieces of the decision-making loop that are delegated to computers, precisely which AI technology is incorporated, and which physical capabilities are granted to the system.
The uncertainty that AI-enabled AWS must be able to handle derives from operating environments that are only partially observable and/or nondeterministic. Programming a computer with vast data sets that can account for every possible outcome to even simple problems is often mathematically infeasible, factually impossible, or undesirable due to other mission constraints or restraints. Thus, AI in the context of AWS must necessarily account for potentially immense uncertainty while achieving a desired end state. Uncertainty in AI manifests itself primarily in two ways: first, it is sometimes infeasible or impossible to establish exceptionless rules for the system to follow; second, the system may be ignorant regarding some aspects of its operating environment.

So how does AI account for this uncertainty? It does so by linking a computer's decisions to the probability of certain outcomes and the utility of such outcomes. “Probability provides a way of summarizing the uncertainty that comes from our laziness and ignorance.” Utility theory establishes “preferences between the different possible outcomes of the various plans.” “The right thing to do--the rational decision--therefore depends on both the relative importance of various goals and the likelihood that, and degree to which, they will be achieved.” An AI is considered to make rational decisions “if and only if it chooses the action that yields the highest expected utility, averaged over all the possible outcomes of the action.” Simply put, humans must give computers decision-making priorities based on levels of certainty. We must program the computer to achieve the “best” possible outcome under the circumstances.

The notion that AI must acknowledge and account for uncertainty in its programming is important in the evaluation of AWS. The sophistication of the programming in its ability to account for uncertainty on the battlefield determines whether the system can comply with IHL. The AI might possess crude and unsophisticated deterministic software, a more sophisticated Bayesian network, or some unforeseen technology that surpasses what we can currently envision. The bottom line is that we must be able to account for uncertainty in the programming of the system such that we can reasonably predict IHL compliance despite the inherent complexities of combat. Depending on the specific characteristics of the system, this prediction could be simple, impossible, or somewhere in between.

### B. Predictability and Uncertainty in AWS: Tracing Decisions to Kill Back to a Human

Robots seem frightening because, despite their potential power, they inherently lack the context from which to discern the appropriate application of force. They do not know anything, except what they are told through programming. Suppose a human being was born at age 21, fully grown and strong but lacking the knowledge that the average adult would have gained during their two decades of life. That person would know nothing about social norms, rules, or consequences. Babies are not threatening because they are powerless to act on their irrational and undeveloped thoughts. A full-grown person with the temperament of a toddler, however, would be very dangerous. Then again, even the “newborn adult” described above could process the emotions of other humans and begin adapting its behavior. The same is not necessarily so for a machine.

But this conundrum is mitigated by the fact that, unlike humans, machines do not possess free will. As discussed in Part II.B.5, they generally do as they are programmed. They handle uncertainty in the way we tell them to handle it. Their learning is bounded by the ways we tell them to learn. We decide if they have a neural network, no network, or a Speak & Spell for a processor. As such, the responsibility lies with those designing AWS to account for uncertainty, both internal to the system and in handling its external environment, in a responsible manner that ensures we can reasonably predict IHL compliance.
With respect to environmental uncertainty, we may safely assume that AWS will not be able to sense and consider all of the variables present on the battlefield that humans might consider. But AI within AWS will not necessarily need to consider all the same factors that a human might in order to be lawful. Instead, AWS may be able to compensate for a lack of situational awareness in certain respects through other capabilities and/or limitations. For example, a system may not need to complete the complex task of discerning hostile intent because it need not act out of self-preservation. If the AWS is either expendable or well hardened, the system could be restricted to responding only to hostile acts, or forbidden entirely from acting in its own defense. The evaluation of uncertainty and how it is handled by the system should not be confused with the ways in which a human would conduct the process.

Our technological inquiry into whether an AWS handles uncertainty acceptably hinges on whether the actions of the system that result in the loss of human life are predictable enough to be traced back to a human decision to attack a target or class of targets. There are infinite ways in which puzzle pieces from the OODA loop could be stitched together to demonstrate that the decision to kill was reasonably made by a human. Removal or insertion of any given piece could make the difference in determining whether or not a human retained control of this decision and a concurrent IHL evaluation. We must therefore establish broad principles that will allow us to avoid the development of unlawful autonomy in weapon systems. A brief review of the basic IHL principles is first required.

IV. Overview of International Humanitarian Law

The principles of IHL guide the conduct of belligerent parties at all times. They constitute the foundation for how military forces prosecute lethal attacks. The principles are general in nature, and they have withstood the test of time and endured through the development of innumerable means and methods of destruction. The IHL principles are: military necessity, distinction, proportionality, and preventing unnecessary suffering.

The principle of military necessity holds that if a target is “indispensable for securing the complete submission of the enemy as soon as possible,” and an attack upon it is not otherwise illegal, then it is a valid target. Targets are persons and objects “which by their nature, location, purpose, or use make an effective contribution to military action” and whose destruction or neutralization “offers a definite military advantage.” Military necessity does not justify targeting something that is otherwise illegal, and it is not a defense to a violation of IHL.

The principle of distinction holds that belligerents may only attack targets that are valid military objectives. As a subset of the distinction principle, the requirement to take precautions in the attack mandates that a belligerent take active steps to determine whether persons and groups are civilians or combatants and to direct operations only against combatants. This is an affirmative duty on the part of the belligerents. Civilians and their property are of course generally protected from attack.

Proportionality under IHL means that the anticipated loss of civilian life and damage to property incidental to attacks must not be “excessive in relation to the concrete and direct military advantage anticipated.” Thus, collateral damage to civilian personnel and property incurred while attacking a military objective may not be disproportionate to the advantage gained. Military advantage is not weighed based on tactical gains alone. The “expected advantage should be seen in relation to the attack as a whole,” and it is linked to the full strategic and operational context of the attack.
With respect to combatants in a conflict, the prohibition against causing unnecessary suffering prohibits the use of weapons that by their nature cause unnecessary suffering and the use of lawful weapons in a manner that is intended to cause unnecessary suffering. There is no simple test to determine whether the use of a weapon would constitute unnecessary suffering. Nonetheless, as a matter of policy, the U.S. Department of Defense (DoD) reviews every weapon in its inventory to ensure that it does not by its nature cause unnecessary suffering or otherwise violate IHL or other laws. Many other countries do the same, either voluntarily or out of treaty obligations. With these IHL tenets in mind, we may now describe principles that inform the responsible development of AWS.

V. Five Principles: Avoiding Unlawful Autonomy

A. Principle 1: The decision to kill may never be functionally delegated to a computer

Machines will not develop human-like cognitive qualities any time soon. As such, the discussion about machine decision-making must be focused on the potential for functional delegation of the decision to kill. The question of whether or not a target or class of targets can be attacked under the given conflict rubric is inherently a human burden. For this reason, humans must retain control over adequate fragments of the OODA loop. We cannot predict the weapons technology that might exist in the future, but we must ensure that within the OODA construct, as applied to targeting decisions, every future weapon system retains sufficient human input such that the decision to kill is not functionally delegated. This is admittedly the most nebulous aspect of these principles. It is also, however, the most significant. If we cannot predict the weapons that will emerge in the future, we must endeavor to describe in more detail what it means to functionally delegate a decision.

The primary inquiry in this context is how confidently we can establish in advance that a weapon system will kill the intended people or classes of people and destroy the intended objects or classes of objects. Targets may be attacked only if they are legitimate military objectives. Whether a target or class of targets are valid military objectives is a decision that must be left to a human. Thus, we must be able to establish through design and testing that an AWS is reasonably expected to attack only those targets or categories of targets that a human has determined to be valid. It is important to note that this principal does not imply that a human must provide input to an AWS that is temporally proximate to lethal kinetic action, a point that will be further explained in subsequent principles.

By way of example, suppose that during an international armed conflict an unmanned submarine is allowed to loiter in international waters and that it is programmed to destroy any and all enemy warships that it identifies. Assume that, through testing, we can ensure to a reasonable certainty that it will attack only these warships and no other kind of ships. After deployment, the submersible spends a great deal of time ignoring ships it cannot yet positively identify. After some time, perhaps months, it locates a convoy of enemy warships and attacks each of them. In this circumstance it would be inaccurate to say that a machine “decided” to attack the enemy ships. Although it enjoyed a significant degree of autonomy, the decision to designate the enemy ships as valid military targets was made by a human during the design and programming of the submersible.

But what if an AWS was granted far broader discretion—if certain pieces of the OODA loop were delegated that obscured our ability to reasonably conclude that a human being made a decision to kill? Suppose that the above hypothetical is modified so that technology has not advanced to the point where we could assure that the submersible would attack only enemy warships. Assume also that the submersible is granted broader discretion over where it may roam the seas. Under these conditions, we could no longer reasonably expect the system to comply with IHL or the law of neutrality.
is not to say, of course, that such a system would be illegal *per se*. But our concerns over whether the system could comply with IHL, and in what circumstances it could be lawfully employed, are greatly increased. This is because we are no longer reasonably certain that the system will attack only valid military objectives. Notably, the submersible is not “deciding” to attack targets; we simply cannot determine which ones it will attack. In this scenario, too many pieces of the OODA loop may have been ceded to the system to be able to state that a human decided to attack a target or class of targets. In this hypothetical, the decision to kill may have been functionally delegated to the AWS.

The inquiry therefore hinges on the combination of capabilities and authorities granted to the AWS. Capabilities are a relatively familiar concept: what platform is the system deployed on and what weapons is it armed with? As described in Part II.B.5, these factors are subordinate in the analysis to the authorities granted through computer programming. Authorities relate to questions such as, is the system deterministic and, as such, inherently more predictable? Or has it been granted learning AI, which may make the inquiry far more complex? Can the learning be bounded in a way that alleviates concerns about IHL compliance? These issues are discussed below in Principle 2.

**B. Principle 2: AWS may be lawfully controlled through programming alone**

Developers of AWS can ensure that these systems comply with IHL through their programming, even in cases where the system is granted advanced learning AI. Compliance with IHL was more straightforward in the case of legacy weapon systems, which employed only simple and deterministic AI. Essentially, the prediction regarding what a machine would do in given circumstances was akin to a flowchart of “if/then” determinations. But future AWS that possess learning capacity may still comply with IHL, depending on which puzzle-shaped pieces of the OODA loop they are granted and, more importantly in the learning context, on how the system is bounded.°

The challenge of this principle lies in the performance of AI in non-deterministic or partially observable environments, the so-called “fog of war.” The dilemma is how to develop algorithms, establish training data,° and equip AWS to learn in a battlefield environment, which is one of the most confusing and chaotic experiences thinkable. Computer scientists currently struggle to assure reasonable predictability for AI in rigidly controlled laboratories.° How can we envision AI that could operate with reasonable predictability in armed conflict? The answer, at least in the near term, lies in what tasks we can reasonably expect machines to perform and which combinations of OODA loop fragments cannot safely be delegated to them.°

It is safe to assume that in the near term AWS will not be able to perceive, process, and act upon all of the factors humans consider before employing lethal force.° However, it does not follow that AWS cannot comply with IHL, even on complex battlefields, while employing learning AI. The inquiry will hinge on the aspects of the uncertain environment for which the AWS must account and, in turn, the range of available responses provided to the system through its authorities and capabilities.

Since AWS will likely not be able to account for all variables and associated uncertainty on a battlefield, we must inquire into which aspects of the environment are particularly relevant to an IHL compliance inquiry. The first step in this process is to catalogue the range of relevant variables we expect the AWS will encounter.° Within this range, the second step is to establish which variables we expect the AWS to be able to observe. The third step is to define which variables within the observable range could bear on IHL compliance. The final step is to determine which of the observable IHL-relevant variables that we expect the AWS to encounter will be affected by learning AI granted to the system. After narrowing our category of inquiry in this manner we may arrive at a much smaller and more manageable set of variables.°
From this point, the AWS can be programmed to evaluate the probability of certain outcomes as compared to the expected utility of particular actions. Perhaps future systems will need to incorporate highly advanced structured representations of the environment, but not necessarily. The key will be to carefully delineate what the AWS can sense, what it must consider, and therefore the level of sophistication required of the onboard AI.

Suppose that facial recognition software and optical technology advances to the point where airborne drones can affirmatively discern the identity of an individual from afar. The aircraft also have improved efficiency in multiple systems that enable them to loiter for months. Such a system hypothetically conducts a grid-style search for a senior leader from a designated-hostile terrorist group with instructions to attack once it obtains positive identification. After searching for nearly a month, the drone locates a man it believes with 49 percent certainty is the terrorist leader; it takes no action and continues to track the man. The system continues to refine its facial recognition analysis so that by the next day, the drone determines it is 97 percent certain that the man identified is the proper target. Having studied the leader's habits, the system already knows that the terrorist leader goes for a walk alone at about 0800 every morning. As the terrorist leader strolls alone the following day, and without any input from a human, the drone fires a low-collateral damage projectile that has no explosive charge. The weapon guides itself precisely into the forehead of the terrorist leader, killing him instantly.

Consider first the distinction evaluation in this hypothetical. The drone located what it suspected was the target but was not confident even to a “more likely than not” standard that it had positive identification. In this case, autonomy decreases the likelihood that an innocent person will be killed. The drone, unconstrained by limited time on station and unaffected by the natural human tendency to jump at the opportunity in spite of uncertainty, simply waited. The bias of humans who must decide whether to act during these windows of opportunity is unarguably in favor of attacking. The drone knows nothing of wanting to “win” or please its boss. It only knows, so to speak, that the criteria for attack have not been satisfied.

Next consider the IHL requirement to take precautions in the attack in order to minimize “incidental loss of civilian life, injury to civilians and damage to civilian objects.” In this hypothetical, the drone dispassionately evaluated the situation until it determined the optimal moment when it could achieve the mission and satisfy these requirements. If on the other hand a human pilot was asked to make a shoot/no-shoot decision, he or she would currently have limited time on station to make this call. The human pilot might not have time to establish pattern of life. A pilot would feel pressure to attack for fear that the window of opportunity might close. As such, a human would be more likely to accept greater potential for civilian harm. Although arguably not as pressing as piloted aircraft, a remotely piloted aircraft is also controlled by a human who feels the same kind of pressure.

In this circumstance, IHL compliance was assured through programming alone. The human emotions, which distract from achieving an optimal result under IHL, were eliminated. The AWS was not called upon to do more than could reasonably be sensed, processed, or acted upon by a machine. Through programming, we may therefore leverage the strengths of machine learning while avoiding the pitfalls associated with trusting machines to complete tasks where their performance would be unacceptably unpredictable.

C. Principle 3: IHL does not require temporally proximate human interaction with an AWS prior to lethal kinetic action

As a direct corollary to Principle 2, there is no requirement based in IHL that a human must interact with an AWS at or near the time lethal action is taken. As a policy matter, human involvement that is proximate to the point of lethal action might be a good idea under some circumstances, and in others it might not. But there is no legal
requirement for it. This flows from the fact described in Principle 2 that AWS are capable of being lawfully controlled through programming alone.

Some argue that having a human involved in the “Act” phase of the lethal decision-making process is always desirable because this will inherently improve the overall performance of weapon systems. This argument fails for three main reasons. First, alluded to previously, is the significant possibility of a future conflict with a peer competitor. Systems that require a human to approve final lethal kinetic actions will likely be incapable of competing at machine speed with sophisticated peer-competitor opponents in the future. Other countries already claim to be developing systems that could potentially take action independent of human approval processes. In order to be postured to meet such threats, the United States and its allies may need to possess systems that can respond at machine speed. Otherwise, enemy systems massed at critical locations on the battlefield could overwhelm the ability of U.S. forces to react.

Second, manned systems are a significant drain on personnel, training, and budget. It takes years of instruction and millions of dollars, for example, to train one human pilot in the U.S. military. Once employed, humans tend to wear out if subjected to constant combat operations, even if operating aircraft remotely. We simply may not be able to keep pace with the production and employment of AWS by potential adversaries if we continue to rely on direct and proximate human involvement in the lethal decision-making cycle. Again, this is highly contextual. The same argument would not be nearly as convincing if the particular AWS was designed for any operations short of international armed conflict (IAC) with a peer competitor. Then again, the idea that conflict status and intensity might inform the employment of certain weapon systems is not a novel concept.

Third, human involvement is not always helpful with respect to IHL compliance. Some groups have called for “meaningful” human control of AWS. This Article does not engage directly with that particular debate. However, the conception of AWS presented here suggests that the mere fact that a human pushed a button to either approve or disapprove lethal kinetic action does not necessarily correlate to improved IHL compliance.

The issue of whether or not a human must be involved just prior to lethal action by an AWS is a hotly contested matter. The answer will depend on the specific design and intended use of the particular AWS contemplated. There is no per se requirement from a legal standpoint that a human be involved at or near the point of lethal kinetic action.

D. Principle 4: Reasonable predictability is required only with respect to IHL compliance, but will hinge on the specific fragments of the OODA loop granted to the AWS

Our ability to predict the actions of future AWS must be based on a reasonableness standard. This is the standard by which we have historically judged weapons systems in determining whether or not they could comply with IHL. But the predictability of the AWS must, from a legal standpoint, be reasonable only as it bears on our ability to understand whether the system will comply with the law. This means that the system may in fact be lawfully unpredictable in certain ways. So long as the ways in which the system is unpredictable are reasonably unlikely to render an AWS action unlawful, the system may be lawful.

That being said, it would be ill advised to assume that simply because an AWS was predictable in the manner by which it “selected and engaged” targets that it would thereby be lawful. Nor in assessing a system's conformity with IHL should we become overly focused on any particular “critical function.” These factors may bear significantly on the analysis of whether a future AWS is per se unlawful or not. They will not, however, be dispositive. Those reviewing future
AWS for compliance with IHL will need to carefully scrutinize which specific pieces of the OODA loop have been granted to the AWS and, in particular, how machine learning is inserted into this process. Two examples illustrate this point.

First, suppose a legacy drone system such as the MQ-9 Reaper was retrofitted with technology that allowed it to use machine learning to select its route to and from the target area, but was otherwise identical to the system as currently fielded. A remotely stationed pilot would still be required to make the final decision on whether to strike a target. As such, the system could not “select and engage” any target without human approval. Also, as currently posited, the drone would not have any autonomy in its “critical functions.” If left unbounded, however, the relatively simple machine learning granted to this system could violate IHL. For example, the system might decide that the quickest way to the battlefield is a straight line--directly through an air traffic control scheme in neutral airspace. Though the remedy to this dilemma is straightforward, the point still stands that potential threats to IHL compliance linger in other combinations of pieces from the OODA loop than simply the moment of kinetic action.

Next, consider the hypothetical unmanned submarine described in the discussion of Principle 1. The system is able to positively identify and attack targets without human intervention. If it is able to positively identify an opposing belligerent's warships, it sinks them without obtaining permission from a human. The submarine is plainly able to “select and engage” targets and has a great deal of autonomy in its “critical functions,” but its IHL compliance is not in dispute. This is, of course, a very narrow hypothetical. If the facts were changed to take the example out of the IAC construct of status-based targeting under the laws of naval warfare, then the evaluation of the system would take on a different form. Clear regulations on the employment of such a system may need to be issued along with its fielding to the military. It provides an example, however, of an AWS with sufficient autonomy to conduct the targeting OODA cycle without temporally proximate human input, but which would nevertheless cause little concern from an IHL compliance standpoint.

*423 Thus, the only aspects of future AWS that must be reasonably predictable are those that bear on their ability to comply with IHL. Which particular pieces of the puzzle these are will vary widely based on the specifics of the system being developed. It is impossible to predict the future and, as such, impossible to delineate ahead of time which aspects of all AWS must be scrutinized in order to ensure compliance with IHL. We must not attempt to create law that solves intractable problems within systems that do not yet exist. In much the same way, we must also temper the temptation to focus on the capabilities of developing systems by also inquiring into the limitations of their capabilities.

E. Principle 5: Limitations imposed on an AWS may compensate for performance shortfalls

While it is true that future AWS may need advanced AI and machine learning in order to match peer-competitors, it does not necessarily follow that this will lead to the development of sentient “killer robots.” To assume so ascribes too much capability from a technological standpoint to systems that are advanced only in narrow, bounded ways. This assumption also fails to recognize that machines will not necessarily be called upon simply to substitute for humans. The focus of the IHL inquiry should therefore delve not only into the capabilities of a given system, but also scrutinize the ways in which the capacity of the system could be bounded through limitations on authorities and capabilities.

By way of example, Watson is an AI that is currently quite adept at defeating human opponents at answering trivia questions during the game show Jeopardy. This seems highly sophisticated until one reflects on all the things that Watson cannot do. To wit, Watson cannot go anywhere or kill anybody. Even if Watson were outfitted with weapons and mobility, it would have no ability to use them because it was never programmed to do so. Simply put, Watson is in this context quite good at only one thing and that is trivia.
At first, this seems somewhat trivial, but the point is by no means glib. Even if we reassure ourselves that the singularity is not near, we should not thereby be satisfied that AWS do not pose a significant threat to IHL compliance. The problem is that some researchers evaluate the technology backwards. Instead of focusing on all the interesting things *Watson* can do, we should instead be asking what it cannot do. For it is only after we satisfy ourselves that a certain system will not be able to complete a certain required task which might affect IHL compliance that we can adequately compensate for the shortfall. Again, the real dilemma with the ability of future AWS to comply with IHL is not in the production of machines that are smarter than us, it lies in the development of *424* systems which are quite smart, but not smart enough. There are three primary ways in which technology can be designed to account for this threat.

First, we must not assume that machines will simply substitute for humans in any mission set. Instead, we must recognize that machines will likely team with humans in ways that leverage the strengths and weaknesses of both. 194 Autonomy may simply augment human actions rather than replace them. 195 There may be situations in which AWS can and should be deployed with the authority and capability to take lethal action without temporally proximate human input, such as the future conflict with a peer competitor described above. This will not always be the case, however, and so we must carefully evaluate the unique aspects of how a machine could team with a human rather than whether it would be able to replace a human. The answer will depend heavily on the mission set.

Second, AWS can be limited by restricting their platforms and available weapons. We might not entrust the future drone hypothesized in the discussion of Principle 2 with a missile like those currently carried by remotely piloted aircraft. We would have concerns in that instance over collateral damage. 196 We may instead require it to employ extremely-low-collateral-damage weapons such as a high velocity non-explosive bolt. But it would be erroneous to conclude that because an advanced AI might be unable to process and consider all the complexities of the battlefield that we could not incorporate advanced AI and machine learning into AWS. The destructive potential of any system can be limited through its physical capabilities.

Third, as described by Principle 2, AWS may be bounded through the authorities that they are granted in their programming. Deterministic systems will remain predictable enough that we can ensure IHL compliance through simple “if/then” type programming. But even more advanced agents may comport with IHL if we are able to reasonably predict how the agent function and program will respond in certain environments. Future AWS may need to incorporate highly complicated factored and structured representations of the environment in order to account for the complexities of the battlefield. This is not to say, however, that an AWS must account for all factors that a human might consider in arriving at a decision to employ lethal force. By carefully delineating the variables that the agent must observe and assess, we can establish ahead of time whether the agent program and function will collectively be reasonably likely to meet rational *425* objective IHL standards. If we cannot predict precisely how the machine will learn, then the agent program could be provided with hard rules to bound its behavior. 197 If the machine was still unpredictable in ways that adversely impact IHL compliance to unacceptable levels, it would need to be re-designed.

**Conclusion**

Autonomy in weapon systems will likely continue to evolve as technology advances. It is incumbent upon those responsible for the legal review, policy, acquisitions, and systems design of future AWS to ensure compliance with IHL. The discussion regarding potential ramifications of increasing autonomy in weapon systems must likewise evolve from theoretical to practical. The principles described in this Article are intended to serve as the foundation for guidance that will help ensure the lawfully responsible development of autonomy in weapon systems.
The U.S. DoD should take the lead in this regard by incorporating practical legal guidance for the responsible development of autonomy in weapon systems into policy. Policies that currently exist should be extended to delineate the specific, substantive areas of focus for those who seek to develop lawful AWS. The principles proposed by this Article seek to reconcile the need for practical guidance with the perils of crafting rules that are either too broad or unduly narrow.

The question of whether creating any particular AWS is a wise policy decision must likewise be carefully scrutinized. There exists a myriad of non-legal concerns that must be addressed. The balance between developing systems that will facilitate the security of free civilization or instead usher in avoidable death and suffering is a delicate one.

Footnotes

1. Judge Advocate, Lieutenant Colonel, U.S. Marine Corps. Associate Director, Stockton Center for the Study of International Law, U.S. Naval War College, and Fellow, Center on National Security and the Law, Georgetown University Law Center. The views herein should not be attributed to any of the author's institutional affiliates, to include the U.S. Department of Defense.


4. Among the legal concerns, those regarding the application of IHL to AWS are arguably most pressing. Within the IHL context, questions arise primarily regarding: (1) how to ensure AWS comply with the general principles of IHL, and (2) how principles of accountability (such as command responsibility) will apply to AWS employment. This Article focuses on the technology of AWS and how it will intertwine with the application of IHL principles to AWS. There are also a multitude of non-legal concerns raised by AWS, including potential ethical dilemmas.

4. Compare, e.g., Bonnie Docherty, Killer robots are ‘quickly moving toward reality’ and humanity only has a YEAR to ban them, expert warns, DAILY MAIL (June 17, 2016), http://www.dailymail.co.uk/sciencetech/article-3647006/Killer-robots-quickly-moving-reality-humanity-YEAR-ban-expert-warns.html (stating erroneously that “[r]emoving humans from the targeting decision would create a dangerous world. Machines would make life-and-death determinations outside of human control. The risk of disproportionate harm or erroneous targeting of civilians would increase. No person could be held responsible.”).
Some government officials in the United States and abroad have privately expressed the view that we should not be overly concerned about AWS because, for a number of policy and practical reasons, governments will not develop unpredictable AWS that violate IHL because it would not be in their self-interest. This argument is unsatisfying in that entrusting weapons policy simply to self-interest invites deviation from policy when interests change instead of discouraging it.


See, e.g., Benjamin Wittes, Lecture at Georgetown Univ. Law Center (Apr. 16, 2013) (arguing that we should not adopt a treaty *ex ante* to outlaw AWS); but see Tom Malinowski, Lecture at Georgetown Univ. Law Center (Apr. 16, 2013) (arguing in support of his Lawfare blog posts criticizing AWS).


The general sense from those charged with developing AWS policy is that there is little practical legal guidance available. See Workshop on Unmanned Systems held by Deputy Assistant Secretary of the Navy for Unmanned Systems & Navy Unmanned Warfare Systems Directorate (OPNAV N99), San Diego, Cal. (July 26-29, 2016).

A failed attempt to outlaw the discharge of any weapons from aircraft is one particularly shortsighted example of this concept. See THE LAWS OF ARMED CONFLICTS: A COLLECTION OF CONVENTIONS, RESOLUTIONS, AND OTHER DOCUMENTS 309 (Dietrich Schindler & Jiri Toman eds., 4th ed. 2004) (describing the process); *see also* Declaration (XIV) Prohibiting the Discharge of Projectiles and Explosives from Balloons, Oct. 18, 1907, 36 Stat. 2439; Declaration (IV, 1) to Prohibit for the Term of Five Years, the Launching of Projectiles and Explosives from Balloons, and other Methods of Similar a Nature, July 29, 1899, 32 Stat. 1839.

*See generally* YORAM DINSTEIN, WAR, AGGRESSION AND SELF-DEFENCE 16 (5th ed. 2012) (discussing the scope of IHL).

*See, e.g.,* COMMENTARY TO GENEVA CONVENTION I FOR THE AMELIORATION OF THE CONDITION OF THE WOUNDED AND SICK IN THE ARMED FORCES IN THE FIELD 39 (Jean Pictet ed., 1952) (Regarding the “origin and development of the idea” for standards of humane treatment embodied in Common Article 3, “[t]he principle of respect for human personality, which is at the root of all the Geneva Conventions, was not a product of the Conventions. It is older than they are and independent of them.”).

These principles are discussed in greater detail in Part IV.

*See 1 CUSTOMARY INTERNATIONAL HUMANITARIAN LAW* ix (Jean-Marie Henckaerts & Louise Doswald-Beck eds., 2005) (“The laws of war were born of confrontation between armed forces on the battlefield.”).
See e.g., Workshop, Syria: Can International Law Cope?, Stockton Ctr. for the Study of Int’l Law, U.S. Naval War College (Nov. 16-18, 2015).

This Article examines the particular problems of such an approach in Part II.B.1. Although the law generally attempts to use historical examples as a framework to approach future dilemmas, this approach proves unsatisfactory in the context of AWS.

For an in-depth discussion of this matter, see generally Christopher M. Ford, Autonomous Weapons and International Law, 67 S.C. L. REV. (forthcoming 2017).

See Phalanx Close-in Weapons System, RAYTHEON.COM, http://www.raytheon.com/capabilities/products/phalanx/ (last visited Mar. 30, 2017) (“At sea, the Phalanx close-in weapon system--a rapid-fire, computer-controlled, radar-guided gun system--is designed to defeat anti-ship missiles and other close-in air and surface threats. The land-based Phalanx weapon system is part of the U.S. Army's Counter Rocket, Artillery and Mortar systems used to detect and destroy incoming rounds in the air before they hit their ground targets.”).


See Learn, supra note 6 (“[F]ully autonomous weapons ... would be able to choose and fire on targets on their own, without any human intervention.”).

In the U.S. federal courts, this concept is reflected in the justiciability doctrine, which holds in part that courts will not rule on cases or controversies that are not “ripe” for consideration. See Erwin Chemerinsky, A Unified Approach to Justiciability, 22 CONN. L. REV. 677, 677 (1990) (“Familiar and well-settled law requires that, in order for a federal court to hear a case, several justiciability doctrines must be met: the case must not present an advisory opinion; there must be standing; the case must be ripe; it must not be moot; and it must not present a political question.”).

The manner in which one balances these important interests often hinges on the institutional biases inherent in one’s profession. Those in the military acquisitions field are justifiably concerned that they obtain the greatest possible advantage over peer competitors. On the other hand, professionals who strive to enhance humanity during the conduct of hostilities rightly focus on alleviating unnecessary death, destruction, and suffering.

“We can only see a short distance ahead, but we can see plenty there that needs to be done.” Alan Turing, Computing Machinery and Intelligence, 49 MIND 460 (1950). An intractable problem is defined as “not easily governed, managed, or directed.” Intractable, MERRIAM-WEBSTER.COM, http://www.merriam-webster.com/dictionary/intractable (last visited Mar. 30, 2017).

See Smithsonian Nat’l Museum of Natural History, Early Stone Age Tools, HUMANORIGINS.SI.EDU, http://humanorigins.si.edu/evidence/behavior/stone-tools/early-stone-age-tools (last visited May 9, 2017) (“The earliest stone toolmaking developed by at least 2.6 million years ago.”); see also 2001: A SPACE ODYSSEY (MGM 1968) (in the opening scenes depicting the “dawn of man,” human predecessors transition from using verbal and physical threats against a hostile tribe to employing improvised weapons against them.).

The machine tractor is one example.
Consider the advent of the machine assembly line.

See WILLIAM H. BOOTHBY, CONFLICT LAW: THE INFLUENCE OF NEW TECHNOLOGY, HUMAN RIGHTS, AND EMERGING ACTORS 4-5 (2014) (“The digital revolution is only a part of the rapid technological transformation of warfare that we have seen over the last couple of decades .... The ways in which war is conducted and the associated technology that is employed are ... continually changing.”); see also JAMES J. BUSUTTIL, NAVAL WEAPONS SYSTEMS AND THE CONTEMPORARY LAW OF WAR 15 (1998) (“The naval mine technology available today ranges from the simple to the esoteric. Unsophisticated mines from before World War I are used alongside microprocessor mines of the computer age.”).

See DEF. SCI. BD., Summer Study on Autonomy 5 (June 9, 2016), https://www.hsdl.org/?abstract&did=794641 (“Advances in AI are making it possible to cede to machines many tasks long regarded as impossible for machines to perform.”).

The Wright brothers' first flight at Kitty Hawk occurred on December 17, 1903. See NOVA: Wright Brothers' Flying Machine (PBS television broadcast Nov. 11, 2003). The first aerial combat victory occurred barely a decade later during World War I, on October 5, 1914. See Tony Reichhardt, The First Aerial Combat Victory, AIR & SPACE (Oct. 4, 2014), http://www.airspacemag.com/daily-planet/first-aerial-combat-victory-180952933/; see also WILLIAM H. BOOTHBY, WEAPONS AND THE LAW OF ARMED CONFLICT 363 (2009) (“If ... technology may represent a significant military advantage to a state or states, the law can only make a difference if the states concerned can be persuaded to forego that advantage, and that may not be easy to achieve.”). On the other hand, a multitude of technologies developed initially for military application later saw implementation for peaceful purposes. See, e.g., Biography of Dr. Wernher von Braun, MARSHALL SPACE FLIGHT CTR. HISTORY OFFICE, https://history.msfc.nasa.gov/vonbraun/bio.html (last visited Apr. 10, 2017) (“[T]he V-2 rocket was the immediate antecedent of those used in space exploration programs in the United States and the Soviet Union.”).


Zuse invented the Z-3 in Germany. See STUART RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 14 (2010).


See, e.g., EX MACHINA (Universal Pictures 2016) (android becomes self-aware and acts based on self-preservation); WARGAMES, (MGM/United Artists 1983) (computer is placed in control of America’s nuclear arsenal and nearly causes World War III); 2001: A SPACE ODYSSEY, supra note 28 (computer kills the crew aboard a spacecraft because it determines that they are a threat).

The decision to take the life of another human being is one of the most difficult quandaries a rational human may ever face because it goes against the nature and upbringing of many cultures. See, e.g., ROMANS 13:9 (“Thou shalt not kill.”). A decision to kill, even when justified under the law, is intensely personal and for most humans fraught with indecision. See DAVE GROSSMAN, ON KILLING 4 (1995) (“[T]here is within most men an intense resistance to killing their fellow man ....”; see also THE THIN RED LINE (Twentieth Century Fox 1998) (During a full in intense fighting, a battle-weary soldier screams, “Who decides who lives? Who decides who dies?”). Because of this conditioning, the idea that a non-human would be delegated the ultimate task of deciding--in the human sense--who lives and dies is too much for many to accept. But machines do not make decisions in the human sense. Thus, the key question from a technological standpoint is at what point has this burden been functionally delegated to a machine?

A more traditional model of proximate cause in this circumstance would be a soldier aiming a rifle at the enemy and firing. The link between the decision to kill and the death of the enemy is obvious. But if a weapon system is granted the ability to learn based on its environment and then to select from amongst a range of potential targets, the proximity of the decision to kill and the act of killing may be diluted.

On the battlefields of even seventy-five years ago, death came soon after the firing of a rifle or lobbing of artillery rounds. Today, death may result minutes or even hours after the launch of a missile. In the future, lethal kinetic effects may follow days, weeks, months, or even years after the deployment of a weapon system with autonomous attributes.

We could avoid functionally delegating the decision to kill, for example, through carefully tailored and tested computer programming. In the alternative, a control tether might suffice.

Importantly, this does not mean that human involvement is temporally proximate to the moment of lethal kinetic action. This point is discussed in more detail in Part V.


The term “machine decision-making” is shorthand for instances in which machines are delegated by human programmers the authority to complete a task given certain inputs, variables, and/or algorithms.


See HEADQUARTERS, U.S. DEPT OF THE ARMY, FM 6-40, TACTICS, TECHNIQUES, AND PROCEDURES FOR THE FIELD ARTILLERY MANUAL CANNON GUNNERY 1-3 (1996) (describing the process by which deflection and elevation data are calculated).
Assuming the other requirements for accurate predicted fire were met. See id. at 1-3.

And those systems that are not under continuous control, such as landmines, are so deterministic that the humans emplacing them can be certain of the result given a particular input.

Relatively few broad generalizations regarding future AWS withstand careful analysis, but one can reasonably conclude that the computers onboard future AWS may need to have more discretion programmed into their computers than traditional computation-focused systems, such as an artillery computer.

I argue in this Article that they should be at the forefront. See also DUSTIN A. LEWIS, GABRIELLA BLUM & NAZ K. MODIRZADEH, HARVARD LAW SCH. PROGRAM ON INT’L LAW AND ARMED CONFLICT, WAR-ALGORITHM ACCOUNTABILITY (2016) (identifying “algorithmically-derived ‘choices’ and ‘decisions’” as a “central concern regarding technical autonomy in war”).

This is an intentional oversimplification of the question at this juncture. “An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators.” RUSSELL & NORVIG, supra note 36, at 34. A simple reflex agent acts based solely on what it is able to perceive according to “if/then” condition-actions rules such as “if car-in-front-is-braking then initiate-breaking.” Id. at 48-50. Technology advanced long ago past the point where actions taken by machines were the relatively predictable result of simple computers with reflex agents operating on “if/then” condition-action rules.

See RUSSELL & NORVIG, supra note 36, at 28 (“What can AI do today? A concise answer is difficult because there are so many activities in so many subfields.”). We should stop trying to draw a clear line between autonomous and automated. This is a futile effort that attempts to paint over infinite shades of gray with a façade of order. It is also likely a quest to know the unknowable. Most importantly, there is no legal tipping point inherent in these descriptions because they are non-linear at best and arbitrary at worst. See Interview with Paul Scharre, supra note 41 (arguing that it is not helpful to delineate overly generalized distinctions between gradations of autonomy). More automation does not always lead to autonomy or to legal challenges, and as such these categorizations are not useful in describing specific combinations of autonomy that are legally problematic.


Attempts to comprehensively describe legal categories of autonomous systems in toto have borne relatively little fruit. The technology is too diverse to categorize succinctly yet comprehensively from a legal perspective. Whether a machine is allowed to “select and engage” a target may be useful in describing a segment of automation we should take a careful look at due to its operational significance, but it is less helpful in defining a category of automation that is legally objectionable. Instead of attempting to describe and formulate specific rules for the entire possible spectrum of autonomy, we should focus on broad governing principles for the responsible development of AWS.


See Scharre & Horowitz, supra note 57, at 8 (describing the construct of human in, on, or out of the loop).


This model is not a perfect fit insofar as it invites us to ascribe a “decision” to a machine, when in most situations this is misleading as the decision to behave in a certain way is, generally speaking, established by programming. Setting aside that issue, however, the OODA Loop is nevertheless a useful way of envisioning the spectrum of autonomy insofar as it is a
generally accepted method of evaluating human decision-making in the employment of weapons, and thus the ways in which computers might supplant certain actions by humans.


66 Facial recognition technology with remarkable accuracy has already been implemented in social media. See, e.g., Haje Jan Kamps, Apple Introduces Facial and Object Recognition for Mobile Photographers, TECHCRUNCH (June 13, 2016), https://techcrunch.com/2016/06/13/apple-image-and-facial-recognition/.

67 See discussion supra note 52.


70 See Interview with Leslie Pack Kaelbling, Learning and Intelligent Systems Group, Computer Science and Artificial Intelligence Laboratory, Mass. Inst. of Tech., Cambridge, Mass. (Sept. 16, 2016) (explaining the critical role of training environments in machine learning); see also interview with Alan C. Schultz, supra note 68 (noting that it may be possible to reverse engineer why a system took a particular course of action); interview with Naval Undersea Warfare Ctr. autonomous systems specialists, Naval Sea Systems Command, U.S. Navy, Newport, R.I. (Oct. 2015) [hereinafter NUWC Interview] (describing the particular difficulties in this context with neural networks).

71 See Conference Paper, William F. Bundy, Future Maritime Forces: Unmanned, Autonomous, and Lethal, EMC Chair Symposium: Maritime Strategy (Mar. 23-24, 2016), https://www.usnwc.edu/Academics/Faculty/Derek-Reveron/Workshops/Maritime-Strategy/working-papers/bundy.aspx (arguing for the development of “fully autonomous” and “intelligent” unmanned maritime systems); see also NUWC Interview, supra note 70 (supporting generally the premise that future systems will include learning technology).


74 See Interview with Leslie Pack Kaelbling, supra note 70 (expressing skepticism that scientists will continue to be able to reverse engineer actions by a learning system in order to establish why it took a particular course of action); see also Will Knight, The Dark Secret at the Heart of AI, MIT TECH. REV., May/June 2017, at 56-60 (describing the difficulties associated with explaining why deep neural networks take certain actions).
Indeed, no one would argue that a Tomahawk cruise missile “decided” to destroy an object when the GPS location of that target was programmed into the weapon by a human. Delays in temporal proximity between programming and a lethal kinetic event may prove legally significant when evaluating compliance with the principle of proportionality, for example, but they do not necessarily mean that a machine has made the decision to kill.

That is to say, an environment that involves random variables--the opposite of which is a deterministic environment, where the operating environment remains stable and outcomes are determined solely by the actions of the machine in that environment. See RUSSELL & NORVIG, supra note 36, at 43 (contrasting stochastic and deterministic operating environments).

Land mines have been used as a means of warfare for many years and are already governed by specific legal regimes. See Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on Their Destruction, Sept. 18, 1997, 2056 U.N.T.S. 211.


See Interview with Mary Wareham and Bonnie Docherty, Human Rights Watch Arms Div., Washington, D.C. (Jan. 29, 2016) (expressing skepticism that legacy weapons systems would be included in a ban).

Nevertheless, someone who accepts the proposition that a legal framework for future AWS need not revisit the legality of legacy systems might still argue that IHL architecture for AWS should include consideration of legacy systems' features. This is problematic because the model would be forced to contemplate the characteristics of legacy systems while abstaining completely from a fresh review of their legality.

For example, the Samsung SGR-1 presents no vexing weapons-review issues because it is reasonably similar to legacy systems. The system is designed to be deployed in static positions along a demilitarized zone through which no persons are lawfully allowed to pass. See SGR-1, supra note 65. It is able to detect a human being and shoot them, although as currently fielded it requires human approval to engage. See id. It can also determine when a person raises their hands in surrender and hold fire. See IFSEC Global, Intelligent Surveillance & Security System Samsung Techwin, YOUTUBE (Mar. 10, 2009), https://www.youtube.com/watch?v=NevCAx6zWNU (showing mock attackers successfully surrender). But this operating environment is highly constrained and the discretion (shoot/do not shoot) provided to the weapon system is likewise narrow. When one compares SGR-1 to a legacy system such as an electric fence, the similarities are striking, assuming of course that trespassers are put on notice of the system's presence. The most pertinent difference weighs in favor of the legality of SGR-1: an electric fence cannot accept surrender. As such, the system is an excellent example of one that varies in degree of automation from legacy systems but not in kind. That being said, if the SGR-1 was granted more advanced autonomy, such as machine learning, and/or provided with broader physical capabilities, an evaluation under IHL could easily become immensely complicated. For example, if the system was modified such that it could roam the countryside, but was not granted more sophisticated sensor and/or computational capabilities or limited by additional parameters, the evaluation of legality would be markedly different. Likewise, if the system were programmed to learn about human behaviors such as feigning surrender the analysis would be more difficult.

To be sure, there may be some basic commonalities in at least the procedural application of weapons reviews between systems that are different in kind from each other--for example, a laser, a tank, and a cyber weapon. The law is of course the same and the requirements to conduct the reviews are identical. But the principles through which we apply the law to the systems may be quite distinct. In other words, the questions we must ask to inform our decision as to the lawfulness of a given system under IHL will be different based on the kind of system we evaluate. In much the same way that cyberwarfare forces us to re-evaluate how IHL applies, we must understand that autonomous systems are different in kind from legacy systems. See
generally TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE (Michael N. Schmitt ed., 2013) (providing a framework for evaluating cyber weapons that were different in kind from legacy weapons).

News headlines concerning AI often lead with a picture from the “Terminator” movie series. See, e.g., Docherty, supra note 4. The Terminator embodies and promotes popular misconceptions about AI. See Ardalan Raghian & Matthew Renda, When Hollywood does AI, it's fun but far fetched, CNET (June 30, 2016), http://www.cnet.com/news/hollywood-ai-artificial-intelligence-fun-but-far-fetched/ (“Least realistic [depiction of AI in a movie]: ‘The Terminator.’ This pop-culture touchstone is universally reviled by the AI community ... [because it commits] the dual sin of overemphasizing the robotics aspect of AI and also vesting AI with human qualities like a hunger for power and an aptitude for murder.”)

See generally Special Report: The Singularity, IEEE SPECTRUM (2016), http://spectrum.ieee.org/static/singularity. Most AI experts do not believe it is reasonably possible that machines will become sentient beings capable of outperforming humans at thinking and behaving in distinctly human ways. See also DARPA Interview, supra note 73; interview with Paul Scharre, supra note 41; interview with Alan C. Schultz, supra note 68; Alfred Nordmann, Singular Simplicity, IEEE SPECTRUM (June 1, 2008), http://spectrum.ieee.org/robotics/robotics-software/singular-simplicity (concluding that “there is nothing wrong with the singular simplicity of the singularitarian myth--unless you have something against sloppy reasoning, wishful thinking, and an invitation to irresponsibility”); see also DAVID A. MINDELL, OUR ROBOTS, OURSELVES 9 (2015) (“[The myth of full autonomy] is “the utopian idea that robots, today or in the future, can operate entirely on their own.”). We shall therefore set aside the possibility that computers might eventually become “self-aware.” This is a fanciful suggestion and there is little debate that machines possessing self-awareness and human-like discretion would be legally problematic.

See John Vogel, Terminator 2: Judgment Day: Plot Summary, INTERNET MOVIE DATABASE, http://www.imdb.com/title/tt0103064/plotsummary?ref_=tt_stry_pl (last visited Apr. 7, 2017) (“Skynet, the 21st century computer waging a losing war on humans sends a ... terminator back in time to destroy the leader of the human resistance while he is still a boy.”).

But even “true believers” in the singularity do not envision it occurring inside of 30 years. Special Report, supra note 85.


We should of course address AWS that are allocated pieces of the OODA loop but whose predictable execution of those tasks degrades performance of the system below that which is accepted by IHL. See Anderson & Waxman, supra note 72 (questioning “whether artificial intelligence and computer analytic systems could ever reach the point of satisfying the fundamental ... legal principles of distinction and proportionality”).


See Interview with Leslie Pack Kaelbling, supra note 70 (expressing skepticism regarding our ability to predict the behavior of systems equipped with machine learning).

See RUSSELL & NORVIG, supra note 36, at 1 (“[AI] attempts not just to understand [how we think] but also to build intelligent entities.”) (emphasis omitted).
The specific focus of any evaluation of AI depends largely on whether we are more concerned with a system's processes or outputs. See RUSSELL & NORVIG, supra note 36, at 27. And, with respect to processes and outputs, we must inquire as to whether we seek to mimic humans or instead to attain an objectively rational standard. See id. at 1 (“[Success can be measured] in terms of fidelity to human performance ... [or instead] against an ideal performance measure, called rationality.”).

Whether or not a computer is able to mimic the way in which the human brain functions or whether its computational processes appear rational are of no concern in the context of AWS.

See Wittes, Does Human Rights Watch Prefer, supra note 2 (calling for “the development and deployment of automated technologies in those instances in which they would perform better than people and not in those instances in which they would make things worse”) (emphasis omitted).

See Wittes, supra note 7 (arguing that we should not adopt a treaty which outlaws AWS in part because humans have a checkered record of complying with IHL, so we should allow for the possible development of machines that are better at this task).


See RUSSELL & NORVIG, supra note 36, at 59.

See id. at 35 (“Mathematically speaking, we say that an agent's behavior is described by the agent function that maps any given percept sequence to an action.”).

See id. at 59 (“The agent program implements the agent function.”).


The Roomba is limited not only by its physical platform but also by its software. It employs deterministic software that controls its actions based on “if/then” rationale. For example, the Roomba's software tells it, in essence, that “if I have already vacuumed a location, I do not vacuum it again” and “if my battery reaches X percent, I return to my charging station.” The vacuum cannot perceive or respond to environments outside of its programming.


See AlphaGo, DEEPMIND.COM, https://deepmind.com/research/alphago/ (last visited Mar. 29, 2017) (“In March 2016 AlphaGo won 4-1 against ... the top Go player in the world over the past decade.”).


Certain games such as chess are relatively complex yet have a finite number of possible moves. This type of game plays to the strengths of an AI because it can evaluate all possible moves and counter-moves. Other games do not have a finite set of actions. For example, in no-limit Texas hold ‘em poker, a participant can bet any amount of money. Still, AI can compensate by learning how to bluff in ways that are unpredictable to a human and thereby prevail. See Avery Thompson, An AI Just Crushed Poker Pros at Texas Hold ‘Em, POPULAR MECHANICS (Jan. 31, 2017), http://www.popularmechanics.com/technology/a24989/ai-wins-texas-hold-em/. But in all of these games, the AI uses experience to refine its actions in order to achieve optimal results. It has no higher order appreciation for the behavior of its opponents outside the narrow parameters of the game itself.
See Losing Humanity, supra note 1, at 30 (concluding, for example, that yet unrealized “fully autonomous weapons ... would appear to be incapable of abiding by the key principles of international humanitarian law”).


AWS that fight at machine speed and can adapt their behavior quicker than either humans or enemy AWS will arguably be at a natural advantage in future conflict.

See RUSSELL & NORVIG, supra note 36, at 55 (“Learning has another advantage ... it allows the agent to operate in initially unknown environments and to become more competent than its initial knowledge alone might allow.”).

See id. (“The performance element ... takes in percepts and decides on actions.”).

Id.

Id. at 57.

The “critic” is a portion of the agent that provides feedback to the learning element. See id. at 55. The system may also employ a “problem generator” in order to devise novel means by which the drone can achieve its goal of remaining undetected. See id. at 56. The “problem generator” is “responsible for suggesting actions that will lead to new and informative experiences.” Id. It forces the agent to consider courses of action that might appear sub-optimal in the near term but are more successful in attaining long-term goals. Id.

See id. at 57 (“In an atomic representation each state of the world is indivisible--it has no internal structure ... [it is a] state of the world ... whose only discernible property is that of being identical to or different from another [state of the world].”).

This means, for example, that the Roomba could consider the presence of guests and the attitude of its owner in determining whether or not it would set about its mission. See id. at 56 (“A factored representation splits up each state into a fixed set of variables or attributes, each of which can have a value.”).

In other words, the Roomba would be able to ascertain whether the presence of the guests affected the mood of the owner and vice versa.

Losing Humanity, supra note 1.

By way of an example from Losing Humanity, soldiers kicking down doors to residential homes must make split-second “shoot/no shoot” decisions based on determinations of hostile intent. A soldier may use lethal force to counter a hostile act, which is a use of force directed at the soldier or his unit and which is already underway. Hostile intent is inherently more subjective than a hostile act because, in essence, one is trying to predict something that is about to happen. With respect to U.S. military forces, hostile intent is defined in relevant part as the “[t]he threat of imminent use of force against ... U.S. forces.” Chairman
See DARPA Interview, supra note 73; interview with Leslie Pack Kaelbling, supra note 70; NUWC Interview, supra note 70; interview with Paul Scharre, supra note 41; interview with Alan C. Schultz, supra note 68.

Unless of course they are programmed to protect themselves. An AWS could be limited by programming, however, only to use lethal force in response to a hostile act, as opposed to attempting to discern intent. Indeed, this would be a more reasonable requirement for a robot because it does not fear death. The primary reason we allow combatants to use deadly force in response to demonstrated hostile intent is to preserve life, and secondarily to preserve military assets. With an autonomous system, the primary concern is alleviated.

See API, supra note 89, at art. 57 (“[T]hose who plan or decide upon an attack shall ... take all feasible precautions in the choice of means and methods of attack with a view to avoiding, and in any event minimizing, incidental loss of civilian life, injury to civilians and damage to civilian objects.”).

This is especially so in more advanced learning systems such as artificial neural networks, which in their most basic sense attempt to mimic the ways in which biological neurons in the human brain function. See A Basic Introduction to Neural Networks, UNIVERSITY OF WISCONSIN-MADISON COMPUTER SCIENCE DEPT, http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html (last visited Dec. 23, 2016); see also Unsupervised Feature Learning and Deep Learning Tutorial: Convolutional Neural Network, STANFORD UNIVERSITY, http://ufldl.stanford.edu/tutorial/supervised/ConvolutionalNeuralNetwork/ (last visited Dec. 23, 2016) (modeled after the visual cortex of animals, these networks are “comprised of one or more convolutional layers ... and then followed by one or more fully connected layers as in a standard multilayer neural network”). Artificial neural networks “remain one of the most popular and effective forms of learning system.” RUSSELL & NORVIG, supra note 36, at 728. Other cutting-edge technologies may further complicate the question of predictability. For example, polymorphic networks modify their own code while keeping algorithms intact. See Glossary: Polymorphic, SYMANTEC, https://www.symantec.com/security_response/glossary/define.jsp?letter=p&word=polymorphic (last visited Dec. 27, 2016) (“Polymorphic malicious code generates functionally equivalent but distinct copies of itself when it replicates, in the hopes that pattern matching security tools won’t be capable of detecting it, as there is little or no stable pattern of code to match against.”). Transfer learning may also create seams in predictability. See, e.g., “DEFCON 24” Convention, Las Vegas, Nev. (Aug. 4-7, 2016) (“Common machine learning algorithms ... traditionally address isolated tasks. Transfer learning attempts to change this by developing methods to transfer knowledge learned in one or more source tasks and use it to improve learning in a related target task.”) (notes on file with author). When machine learning is exposed to adversarial environments, additional issues arise with respect to predictable performance. See Pavel Laskov & Richard Lippmann, Machine Learning in Adversarial Environments, MIT LINCOLN LABORATORY (June 28, 2010), http://llwww.ll.mit.edu/mission/cybersec/publications/publication-files/full_papers/2010_10_25_Lippmann_MLJ_FP.pdf (noting that when machine learning is employed in hostile environments, “adversaries consciously act to limit or prevent accurate performance”).

RUSSELL & NORVIG, supra note 36, at 693.

See Alan L. Schuller, Focusing the debate on autonomous weapon systems: A new approach to linking technology and IHL, in Meeting Report, Int’l Comm. of the Red Cross, supra note 61, at 27.

See CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at 558 (suggesting that commanders are responsible for war crimes committed by subordinates if they do not take “necessary and reasonable measures in their power to prevent their commission”); see also id. at 563 (“[N]ecessary and reasonable measures” are “limited to such measures as are within someone's power, as no one can be obliged to perform the impossible.”) (citing Prosecutor v. Delalié et al., Case No. IT-96-21-T, Judgment, ¶ 395 (Int’l Crim. Trib. for the former Yugoslavia Nov. 16, 1998)); see also API, supra note 89, at art. 57(4) (“In the conduct of military operations at sea or in the air, each Party to the conflict shall ... take all reasonable precautions to avoid losses of civilian lives and damage to civilian objects.”); Prosecutor v. Stanislav Galic, Case No. IT-98-29-T, Judgment and Opinion, ¶ 58 (Int’l Crim. Trib. for the former Yugoslavia Dec. 5, 2003) (“In determining whether an attack...
was proportionate it is necessary to examine whether a reasonably well-informed person in the circumstances of the actual perpetrator, making reasonable use of the information available to him or her, could have expected excessive civilian casualties to result from the attack.

For example, states may require predictability to a “near certainty.” See generally THE WHITE HOUSE, REPORT ON THE LEGAL AND POLICY FRAMEWORKS GUIDING THE UNITED STATES’ USE OF MILITARY FORCE AND RELATED NATIONAL SECURITY OPERATIONS 25 (2016), https://www.justsecurity.org/wp-content/uploads/2016/12/framework.Report_Final.pdf (explaining that, in the context of targeting, “the United States must have ‘near certainty’ that the terrorist target is present and that non-combatants will not be injured or killed” before lethal action is permitted).

To be clear, this Article does not argue that a human must be “in the loop” and either affirmatively approve or choose not to override the system's decision just prior to lethal kinetic action. Contra Malinowski, supra at note 7 (arguing that while there may be advantages to automation, humans should always be in the decision-making loop in order to approve or deny lethal decisions).

See supra Part II.A.

See supra Part II.B.3.

One of the challenges AI developers face is establishing training and test data that will adequately replicate the experiences faced by an agent outside of the laboratory. This is true even in comparatively simple applications, such as driverless cars. See Interview with Leslie Pack Kaelbling, supra note 70 (describing the challenges). It is difficult to overstate the complexities of a traditional battlefield, and even more so when contemplating future conflict areas.

See RUSSELL & NORVIG, supra note 36, at 480 (“When interpreting partial sensor information, a logical agent must consider every logically possible explanation for the observations, no matter how unlikely. This leads to impossibly large and complex belief-state representations.”).

See Losing Humanity, supra note 1; see also Advancing the Debate, supra note 1.

See Interview with Alan C. Schultz, supra note 68 (taking the position that machines will more likely team with humans on future battlefields than replace them). For example, if we imagine robots in the future playing a game of soccer, most people conjure images of what amounts to metal humanoids on a traditional pitch. But robots might fare better on a hard surface rather than a grass field. Also, a lower center of gravity than a human could be advantageous. Instead of legs, the robots could use flappers like a pinball machine. They might be painted instead of wearing uniforms. The ball might be hard instead of soft. If the machines were programmed properly, referees might not be needed to adjudge penalties. In short, machines will continue to change the way “the game” is played.

See Interview with Alan C. Schultz, supra note 68 (supporting this proposition).

These general concepts are explored via hypotheticals in Part V, infra.

RUSSELL & NORVIG, supra note 36, at 480.

See id.

The context in which AWS are likely to operate amplifies uncertainty for a variety of reasons: Trying to use logic to cope with [complex domains] ... fails ... [due to] ... Laziness: It is too much work to list the complete set of antecedents or consequents needed to ensure an exceptionless rule and too hard to use such rules; Theoretical ignorance: ... science has no complete theory for the domain; Practical ignorance: Even if we know all the rules, we might be uncertain [about a given circumstance due to our inability to sense or account for all variables].

Id. at 481.
Id. at 482. Laziness is the inability to define an exceptionless rule. Ignorance is the inability to know all particular applications of a rule. Id.

Id.

Id.

Id. at 483. “Expected” refers to the average or statistical mean of the outcomes, weighted by the probability of the outcome. Id.

Bayesian means “[b]eing, relating to, or involving statistical methods that assign probabilities or distributions to events (as rain tomorrow) or parameters (as a population mean) based on experience or best guesses before experimentation and data collection and that apply Bayes’ theorem to revise the probabilities and distributions after obtaining experimental data.” Bayesian, MERRIAM-WEBSTER.COM, http://www.merriam-webster.com/dictionary/Bayesian (last visited Dec. 22, 2016).


See DARPA Interview, supra note 73; interview with Paul Scharre, supra note 41; interview with Alan C. Schultz, supra note 68; interview with Lincoln Laboratories autonomous systems specialists, Mass. Inst. of Tech., Lexington, Mass. (Sept. 16, 2016) [hereinafter Lincoln Laboratories Interview].

See discussion supra note 120.

With, for example, protective armor.

Again, the decision need not occur in temporal proximity to lethal action by an AWS.

The law of war principles have attained the status of customary international law. See generally CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17 (describing the customary IHL principles); see also GARY D. SOLIS, THE LAW OF ARMED CONFLICT 250 (Cambridge Univ. Press 2010) (“[T]he core LOAC/IHL principles ... bind every armed force.”).

See, e.g., U.S. Dep’t of Def. Directive 2311.01E, DoD Law of War Program ¶ 4.1 (May 9, 2006, incorporating Change 1, Nov. 15, 2010) (“Members of the DoD Components comply with the law of war during all armed conflicts, however such conflicts are characterized, and in all other military operations.”); see also U.S. DEPT OF ARMY, FIELD MANUAL 27-10, THE LAW OF LAND WARFARE app. A-1 (1956) (IHL seeks to “diminish the evils of war by: a. Protecting both combatants and noncombatants from unnecessary suffering; b. Safeguarding certain fundamental human rights of persons who fall into the hands of the enemy, particularly prisoners of war, the wounded and sick, and civilians; and c. Facilitating the restoration of peace”).

See SOLIS, supra note 150, at 250 (“Despite the codification of much customary law into treaty form during the last one hundred years, four fundamental principles still underlie the law of armed conflict.”) (internal citations omitted).

FIELD MANUAL 27-10, supra note 150, at 4.

Attacks that would produce a “concrete and direct military advantage,” and are not otherwise unlawful, are not prohibited. See API, supra note 89, at art. 51(5)(b); CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 14.

API, supra note 89, at art. 52(2); see also CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 8.

See MARCO SASSÒLI ET AL., HOW DOES LAW PROTECT IN WAR? 387 (3d ed. 2014) (“Necessity ... and self-defence are not circumstances precluding the wrongfulness of IHL violations.”); see also THE U.S. ARMY JUDGE ADVOCATE'S
LEGAL CENTER AND SCHOOL, OPERATIONAL LAW HANDBOOK 11 (David H. Lee ed., 2015) [hereinafter OPERATIONAL LAW HANDBOOK] (noting that military necessity is “not a criminal defense” and does not justify otherwise unlawful acts, but rather “must be applied in conjunction with other LOAC principles”). As such, it is more helpful to think of military necessity as a restraint rather than a permissive concept. In other words, because of the principle of military necessity, a target may not be attacked unless there is a concrete military reason to do so. And, even if there is a military reason to attack it, the other IHL requirements must still be complied with.

See API, supra note 89, at art. 52(2) (“Attacks shall be limited strictly to military objectives.”); CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 7; see also SOLIS, supra note 150, at 258 (noting that the distinction principle is considered customary international law).

See API, supra note 89, at art. 57; CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 15-21; see also OPERATIONAL LAW HANDBOOK, supra note 156, at 12 (“[P]arties to a conflict must direct their operations only against combatants and military objectives.”).

See API, supra note 89, at art. 51-52. However, there are some exceptions to this rule. For example, civilians may lose their protected status based on their actions, such as taking direct part in hostilities, in which case they may be targeted for such time as they do so. See id. at art. 51(3); see also CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 6.

API, supra note 89, at art. 51(5)(b); see also CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 14.

See Rome Statute of the International Criminal Court art. 8(2)(b)(iv), July 17, 1998, 2187 U.N.T.S. 90 (“Intentionally launching an attack in the knowledge that such attack will cause incidental loss of life or injury to civilians or damage to civilian objects ... which would be clearly excessive in relation to the concrete and direct overall military advantage anticipated” is a war crime under the statute.). During American military operations under offensive rules of engagement, the collateral damage estimation (CDE) methodology ensures that attacks conform with this principle. See generally Chairman of the Joint Chiefs of Staff Instruction 3160.01, No-Strike and the Collateral Damage Estimation Methodology (Feb. 13, 2009) (establishing the CDE process). Proportionality in the IHL context, as opposed to the self-defense context of proportionality of force, does not mean that an attacker must use the same type of weapons or force as the enemy. If a target is a lawful one, it may be attacked with any weapon in the military inventory, provided the attack is otherwise lawful. See SOLIS, supra note 150, at 280 (describing common misunderstandings about proportionality).


See API, supra note 89, at art. 35(2)-(3); Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, 1996 I.C.J. 226, ¶ 78 (July 8); CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, supra note 17, at R. 70.

See SOLIS, supra note 149, at 271-72.


See API, supra note 89, at art. 36.

The enemy warships are targetable in international waters at any time due to their status as enemy combatants under the Law of Naval Warfare. See U.S. NAVY, U.S. MARINE CORPS & U.S. COAST GUARD, THE COMMANDER'S HANDBOOK ON THE LAW OF NAVAL OPERATIONS NWP 1-14M/MCWP 5-12/COMDT PUB PS800.7A, para. 8.6.1 (2007) (“Enemy warships ... are subject to attack, destruction, or capture anywhere beyond neutral territory.”).

the Sea, Dec. 10, 1982, 1833 U.N.T.S. 397, art. 2 (“The sovereignty of a coastal State extends, beyond its land territory and internal waters ... to an adjacent belt of sea, described as the territorial sea.”).

168 A simple example is a system whose learning capacity is limited to actions that do not bear on IHL compliance, such as the route an unmanned aircraft will take to the objective based on local weather, assuming air traffic control schemes are accounted for. The more complicated but still tractable problem relates to systems where a system's learning capacity bears directly on lethal actions, such as an unmanned aircraft that uses data from previous strikes to refine its future strike criterion.


170 See Interview with Leslie Pack Kaelbling, supra note 70.

171 See discussion supra note 127.

172 Again, the critical failure of Losing Humanity was that it assumed AWS would necessarily be called upon to perform all of the functions a human would in a time-compressed lethal decision-making cycle.

173 This range may be unwieldy, and if those evaluating the system do not carefully define the relevant variables, the range could become intractable.

174 Care must be taken, however, to map with particularity the ways in which these subsets of variables interact and therefore gain relevance, when they might on their face have none.


176 API, supra note 89, at art. 57.

177 Based on extensive personal experience and interactions with combat aircrew by the author.

178 For example, an MQ-9 Reaper. See MQ-9 Reaper Fact Sheet, supra note 64.

179 A detailed discussion regarding whether or not humans should interact with a given system at the time of lethal action is beyond the scope of this Article. From a technological standpoint, however, it is likely unwise to try answering this question conclusively when uncertainty exists about what AWS may or may not exist in the future. Each new system must be carefully reviewed to ensure it complies with IHL. The question of whether or not humans will be required to interact with the system will depend on the capabilities and authorities granted to that particular AWS.

180 For example, when involvement by a human would not negatively impact either mission accomplishment or legal compliance to an unacceptable level, such as during deliberate targeting of an objective whose location is fixed.

181 One example is based on operational necessity, such as when fighting an international armed conflict against a peer competitor in which AWS must engage enemy AWS at such speed and in such numbers that humans would be unable to effectively react. Another example is based on legal compliance, such as when the involvement of a human (who may act based out of fear, anger, or self-preservation) would reasonably be expected to degrade conformity with IHL to an unacceptable level.

182 See Malinowski, supra note 7 (arguing that a human should remain “in the loop” for any AWS).

See U.S. GOV'T ACCOUNTABILITY OFF., GAO/NSIAD-99-211, MILITARY PERSONNEL: ACTIONS NEEDED TO BETTER DEFINE PILOT REQUIREMENTS AND PROMOTE RETENTION 18 (1999), http://www.gao.gov/archive/1999/ns99211.pdf (“T]he cost to train each military pilot through basic flight training is about $1 million, and the cost to fully train a pilot ... can be more than $9 million.”).


See, e.g., Open Letter, supra note 39 (arguing for “a ban on offensive autonomous weapons beyond meaningful human control”).

See supra Part II.C.

Meeting Report, Int'l Comm. of the Red Cross, supra note 60, at 8.

Id.

See discussion supra note 64.

See Watson, supra note 104.

See Markoff, supra note 104.

It must be noted that Watson technology is being expanded to uses in other industries, but the analogy still applies because the computer will continue to be bounded in innumerable ways.

See Interview with Alan C. Schultz, supra note 68; see also Def. Sci. Bd., supra note 32, at 17.

One excellent example of this is autopilot systems in military and passenger aircraft. These systems do not replace pilots, but they do an excellent job of compensating for areas in which humans are known to perform poorly, such as recognizing the flight profile of the plane and correlating this with required control inputs during emergency situations. See Interview with Alan C. Schultz, supra note 68.


For example, if we are not sure what the system will do in all situations, we could instruct it that, “whatever you do, do not do X.” While it is relatively simple to build basic constraints, more complicated and subjective restraints may prove far more vexing. See Interview with Leslie Pack Kaelbling, supra note 70. This is especially so in complex operating environments. See Lincoln Laboratories Interview, supra note 145. The complexity of the programming dilemma could be reduced in some circumstances by greatly limiting the available options of the AWS. See Interview with Ronald C. Arkin, Associate Dean, Coll. of Computing, Georgia Inst. of Tech., Newport, R.I. (Sept. 22, 2015).

8 HVNSJ 379
An Introduction to
AUTONOMY in
WEAPON SYSTEMS

By: Paul Scharre and Michael C. Horowitz

ABOUT CNAS WORKING PAPERS:
Working Papers are designed to enable CNAS analysts to either engage a broader community-of-interest by disseminating preliminary research findings and policy ideas in advance of a project’s final report, or to highlight the work of an ongoing project that has the potential to make an immediate impact on a critical and time-sensitive issue.
PREFACE

Information technology is driving rapid increases in the autonomous capabilities of unmanned systems, from self-driving cars to factory robots, and increasingly autonomous unmanned systems will play a significant role in future conflicts as well. “Drones” have garnered headline attention because of the manner of their use, but drones are in fact remotely piloted by a human, with relatively little automation and with a person in control of any weapons use at all times. As future military systems incorporate greater autonomy, however, the way in which that autonomy is incorporated into weapon systems will raise challenging legal, moral, ethical, policy and strategic stability issues.

Nation-states and activists in the United States and abroad are already debating how advances in autonomy will influence warfare – and what, if anything, should be done. Activists from 54 non-governmental organizations have launched a “Campaign to Stop Killer Robots.” In May of 2014, state parties to the United Nations Convention on Certain Conventional Weapons (CCW) held initial discussions on autonomous weapons, and discussions will continue at the April 2015 meeting of the CCW.1

Governments and militaries are only beginning to grapple with how to address the challenges and opportunities associated with increased autonomy. Technology is moving fast in this area. Yet few states have guidelines on how autonomy should be included in future weapon systems, with the United States being a notable exception.2

The Center for a New American Security (CNAS) has launched a project on Ethical Autonomy, which will examine the legal, moral, ethical, policy and strategic stability dimensions of increased autonomy in future weapon systems. The goal of CNAS’ Ethical Autonomy project is to help states, activists, academics and militaries grapple with the challenging issues raised by autonomy in future weapons. This dialogue is necessary to ensure an appropriate balance between ethical and strategic stability considerations, technological opportunities and future warfighting needs.

CNAS’ Ethical Autonomy project will result in a series of working papers examining various issues associated with autonomy in weapons, which will ultimately culminate in a final report in late 2015. This first working paper will revolve around framing the issue of autonomy in weapons and providing historical context on the use of autonomy to date. This paper aims to both introduce the topic and clarify the existing debate to help achieve a common understanding of the technology. Subsequent papers will delve deeper into specific legal, moral, ethical, policy or strategic stability issues raised by autonomous weapon systems.

I. INTRODUCTION

The past several years have seen an explosion in commentary and scholarship on the topic of autonomous weapons. The prospect that increased autonomy might change or diminish the control of humans over the use of force has garnered the attention of activists, lawyers, political scientists, ethicists, philosophers, and military and national security professionals, rapidly generating a growing body of work.

It is crucial to understand how today’s weapons already incorporate autonomy. Arguments about whether to incorporate autonomy into future weapons are sometimes made from the perspective of an idealized version of human control, divorced from the realities of war and how existing weapons use forms of autonomy.

In fact, autonomy is already used for a wide variety of military tasks, including many related to the use of force. These include: identifying, tracking, prioritizing and cueing targets; deciding the timing of when to fire a weapon; maneuvering and homing in on targets; and detonation timing. According to research conducted for this paper, at least 30 countries have defensive systems with human-supervised autonomous modes that are used to defend military bases and vehicles from short-warning attacks, where the time of engagement would be too short for a human to respond. And, in a few rare exceptions, autonomy is used for weapons to select and engage targets on their own.

Moreover, there is no internationally agreed-upon definition of what constitutes an “autonomous weapon,” making clear communication on the topic more difficult. The United States Department of Defense, the International Committee of the Red Cross, and the UN Special Rapporteur for extrajudicial, summary or arbitrary executions all use similar definitions, but there is no standard definition that is universally embraced. Activists campaigning to ban “killer robots” (their term for autonomous weapon systems) have yet to put forward a clear definition of their own or even clarify what, precisely, they are advocating should be banned.

This lack of clarity in terminology is further compounded by the fact that some are calling for autonomous weapons to be regulated or banned even before consensus exists on how to define the category. Thus, at present, definitions are tied up in debates over the technology itself. While some degree of debate over the scope of a weapon’s category is intrinsic to negotiations on a potential agreement concerning that weapon, lack of clarity on basic terminology itself is a recipe for disaster.

3. It is most appropriate to think of “weapon systems” that comprise not only the munition itself, but also the launching platform, sensor, targeting process and any other elements such as communications links that are necessary for an engagement to occur, some of which may be distributed across multiple physical platforms. Thus, the more technically accurate term is “autonomous weapon system.” For much of this paper, however, for the sake of brevity we will use the shorthand “autonomous weapon.” In this paper, “autonomous weapon” and “autonomous weapon system” should be treated as interchangeable, and it should be understood that we are referring to not only the munition but the entirety of the weapon system.


5. See Appendix B.

6. See Appendix A.

7. The official page for the Campaign to Stop Killer Robots is available at: www.stopkillerrobots.org/.

This is particularly difficult since there are few, if any, obvious examples of “autonomous weapons.” Activists have cited some weapons today as “precursors” to potential autonomous weapons in the future, but concerns about autonomous weapons are usually aimed at potential future weapons and the general trend of increasing automation, not at any specific weapons that already exist. Thus, one starting point to gain purchase on the definition of autonomous weapons is the understanding that, whatever they are defined to be, they generally do not exist today. Therefore, while there may be rare exceptions, defining “autonomous weapons” in a way that captures wide swathes of existing weapons is almost certainly wrong, since it misses the essence of what future, increased autonomy might bring.

In the interests of helping to clarify the debate, this paper addresses three key questions:

- What is autonomy?
- How is autonomy used in weapons today?
- What is an “autonomous weapon”?

This paper does not examine in detail the legal, moral, ethical and policy issues raised by increased autonomy in future weapons. Rather, the purpose of this paper is to help clarify, as a prerequisite to examining legal, moral, ethical and policy issues, what an autonomous weapon is, how autonomy is already used, and what might be different about increased autonomy in the future.

The intent is that this paper will not only help lay the groundwork for follow-on research as part of CNAS’ Ethical Autonomy project, but also serve as a launching point for others writing on the topic of autonomous weapons. In the span of only a few years, there has been an explosion of interest and concern on the topic of autonomous weapons, and at this juncture taking a step back to review and frame the issue can help focus future debate. The rapid march of technological progress in autonomy, driven largely by civilian advances in robotics and computer processing, underscores a sense of urgency in understanding more clearly the challenges posed by potential future autonomous weapons.

II. WHAT IS “AUTONOMY”?

Different people use the word “autonomy” in different ways, making communicating about trends involving autonomy in weapons particularly challenging. Even setting aside the notion of weapons for a moment, the term “autonomous robot” conjures up wildly different images, ranging from a household Roomba to the sci-fi Terminator. Writers or presenters on the topic of autonomy often articulate “levels of autonomy,” but their levels rarely match up, leading a recent U.S. Defense Science Board report on autonomy to throw out the concept of “levels” of autonomy altogether.  

However one defines autonomy, interest in it is growing among militaries around the world.\textsuperscript{12} In the fall of 2014, U.S. Undersecretary of Defense for Acquisition, Technology and Logistics, Frank Kendall, announced the commissioning of a new study focused on “the science, engineering, and policy problems that must be solved to permit greater operational use of autonomy across all warfighting domains.”\textsuperscript{13} NATO is attempting to better understand the military utility of autonomous systems.\textsuperscript{14} South Korea has developed robotic sentries with “automatic surveillance” to monitor the demilitarized zone with North Korea.\textsuperscript{15} And Russia is developing ground sentry robots to guard missile sites, although their degree of autonomy is unclear.\textsuperscript{16}

In its simplest form, autonomy is the ability of a machine to perform a task without human input. Thus, an “autonomous system” is a machine, whether hardware or software, that, once activated, performs some task or function on its own.

Autonomous systems are not limited to robots or uninhabited vehicles. In fact, autonomous, or automated, functions are included on equipment that people use every day. Most cars today include anti-lock brakes, traction and stability control, power steering, emergency seat belt retractors and air bags. Higher-end cars may include intelligent cruise control, automatic lane keeping, collision avoidance and automatic parking. For military aircraft, automatic ground collision avoidance systems (auto-GCAS) can take control of a human-piloted aircraft if a pilot becomes disoriented and is about to fly into terrain. And modern commercial airliners have a high degree of automation available throughout every phase of a flight. Increased automation or autonomy can have many advantages, including increased safety and reliability, improved reaction time and performance, reduced personnel burden with associated cost savings, and the ability to continue operations in communications-degraded or -denied environments.

Parsing out how much autonomy a system has is important for understanding the challenges and opportunities associated with increasing autonomy. There is a wide gap, of course, between a Roomba and a Terminator. Rather than search in vain for a unified framework of “levels of autonomy,” a more fruitful direction is to think of autonomy as having three main axes, or dimensions, along which a system can vary. These dimensions are independent, and so autonomy does not exist on merely one spectrum, but on three spectrums simultaneously.

The Three Dimensions of Autonomy
What makes understanding autonomy so difficult is that autonomy can refer to at least three completely different concepts:

- The human-machine command-and-control relationship


• The complexity of the machine
• The type of decision being automated

These are all important features of autonomous systems, but they are different ideas, and often people tend to mix them together.

THE HUMAN-MACHINE COMMAND-AND-CONTROL RELATIONSHIP

The first way in which the word autonomy is used refers to the relationship between the person and the machine. Machines that perform a function for some period of time, then stop and wait for human input before continuing, are often referred to as “semautonomous” or as having a “human in the loop.” Machines that can perform a function entirely on their own but have a human in a monitoring role, with the ability to intervene if the machine fails or malfunctions, are often referred to as “human-supervised autonomous” or “human on the loop.” Machines that can perform a function entirely on their own with humans unable to intervene are often referred to as “fully autonomous” or “human out of the loop.” In this sense, “autonomy” is not about the intelligence of the machine, but rather its relationship to a human controller.

THE COMPLEXITY OF THE MACHINE

The word “autonomy” can also be used in a completely different way to refer to the complexity of the system. Regardless of the human-machine command-and-control relationship, words such as “automatic,” “automated” and “autonomous” are often used to refer to a spectrum of complexity of machines. The term “automatic” is often used to refer to systems that have very simple, mechanical responses to environmental input. Examples of “automatic” devices in the civilian world include toasters and mechanical thermostats. In the military world, trip wires and mines can be described as “automatic” devices. The term “automated” is often used to refer to more complex, rule-based systems. Self-driving cars and modern programmable thermostats are examples of such systems. Sometimes the word “autonomous” is reserved for machines that execute some kind of self-direction, self-learning or emergent behavior that is not directly predictable from an inspection of its code. An example would be a self-learning robot that taught itself how to walk or the Nest “learning thermostat.”

Others will reserve the word autonomous only for entities that have intelligence and free will, but these concepts hardly add clarity. “Artificial intelligence” is a loaded term that can refer to a wide range of systems, from those that exhibit near-human or super-human intelligence in a narrow domain, such as playing chess (Deep Blue), playing Jeopardy (Watson) or programming subway repair schedules, to potential future systems that might have human or super-human general intelligence. But whether general intelligence leads to free will, or whether humans even have free will, is itself debated.

What is particularly challenging is that there are no clear boundaries between these degrees of complexity, from “automatic” to “automated” to “autonomous” to “intelligent,” and different people may disagree on what to call any given system.

**TYPE OF FUNCTION BEING AUTOMATED**

Ultimately, it is meaningless to refer to a machine as “autonomous” or “semi-autonomous” without specifying the task or function being automated. Different decisions have different levels of complexity and risk. A mine and a toaster have radically different levels of risk, even though both have humans “out of the loop” once activated and both use very simple mechanical switches. The task being automated, however, is much different. Any given machine might have humans in complete control of some tasks and might autonomously perform others. For example, an “autonomous car” may drive from point A to point B on its own, but a person still chooses the final destination, and potentially even the route. In that case, the “autonomous car” is only autonomous with respect to some functions.

Even in tasks relating to the use of military force, there are different engagement-related functions, not all of which are equally significant when it comes to thinking about the role of human control. Engagement-related tasks include, but are not limited to, acquiring, tracking, identifying and cueing potential targets, aiming weapons, selecting specific targets for engagement, prioritizing targets to be engaged, timing of when to fire, maneuvering and homing in on targets, and the detonation itself.

**Parsing Autonomy**

There is not a single spectrum along which autonomy moves. The paradigm of human versus machine is a common science fiction meme, but a better framework would be to ask which tasks a person does, and which by a machine. A recent NATO policy document came to a similar conclusion, recommending a framework of thinking about “autonomous functions” of systems, rather than characterizing an entire vehicle or system as “autonomous.”

Importantly, these three dimensions of autonomy are independent. Thus, the intelligence, or complexity, of the machine is a separate concept from its ability to perform a task on its own. Increased intelligence or more sophisticated machine reasoning to perform a task does not necessarily equate to transferring control over a wider range of tasks from the human to the machine. Similarly, the human-machine command-and-control relationship differs from the complexity of the machine or the tasks being performed. A thermostat functions on its own without any human supervision or intervention when you leave your house, but it still has a limited set of functions it can perform.

Therefore, when talking about “autonomous weapons,” we must ask: “Autonomous” with respect to which functions or tasks? Which tasks or functions are the critical ones that, when performed by a machine instead of a human, make it an “autonomous weapon?”

First, however, we must start with an understanding of how autonomy is already used in weapons today. If some existing weapon systems are precursors, but autonomous weapons generally do not exist today, then a better understanding of how exactly autonomy is already used – and how it has generally not been used to-date – can help us gain purchase on the question of how to define an autonomous weapon.

---

III. HOW IS AUTONOMY USED IN WEAPONS?

Various forms of autonomy have been used in military systems for over seventy years. The first homing munitions date back to World War II, and automated defensive systems with human-supervised modes have been in existence for decades and are used by many militaries around the globe. Automation is not merely important, but essential for modern militaries to conduct many tasks, including identifying targets by radar or delivering precision-guided weapons.

Since nearly all of the literature on autonomous weapons focuses on the task of selecting and engaging specific targets, the next section delineates how autonomy is used in weapons today along a similar typology. Specifically, we will divide weapons into three categories:

1. Human “in the loop” for selecting and engaging specific targets – Weapon systems that use autonomy to engage individual targets or specific groups of targets that a human has decided are to be engaged.

2. Human “on the loop” for selecting and engaging specific targets – Weapon systems that use autonomy to select and engage targets where a human has not decided those specific targets are to be engaged, but human controllers can monitor the weapon system’s performance and intervene to halt its operation if necessary.

3. Human “out of the loop” for selecting and engaging specific targets – Weapon systems that use autonomy to select and engage targets where a human has not decided those specific targets are to be engaged, and human controllers cannot monitor the weapon system’s performance and intervene to halt its operation if necessary.

Human “in the loop” for selecting and engaging specific targets

World War II saw the advent of homing munitions that used automation to guide projectiles onto their targets. One of the first, the German Zaunkönig (Wren) torpedo, had a passive acoustic homing seeker. Once launched, it would run straight for 400 meters, and then activate its homing seeker, which consisted of two hydrophone receivers. These receivers would sense the noise of a ship’s propeller in the water and then steer the torpedo toward the ship, ultimately destroying it. The Wren was the first of an entirely new class of weapon: the guided munition.

GUIDED MUNITIONS

Guided munitions are “projectiles, bombs, missiles, torpedoes and other weapons that can actively correct for initial-aiming or subsequent errors by homing on their targets or aim-points after being fired, released, or launched.” There are many types of guided munitions, but what they all have in common is that autonomy assists the weapon in striking an individual target or a specific group of targets that a person has decided should be engaged. The person launching the weapon knows what specific targets are to be engaged, and has made a conscious decision that those targets should be destroyed.

---

The Wren acoustic homing torpedo was the first guided munition, but all of the major powers in World War II conducted research on guided munitions, and over subsequent decades they expanded into other domains, including air and ground warfare, and matured considerably. Today, modern militaries around the globe employ guided weapons in the form of air-to-air homing missiles, anti-ship missiles, cruise missiles or air-to-ground guided missiles. The guidance mechanisms of these projectiles vary significantly, but broadly speaking they can be broken into two categories: projectiles designed to hit a particular target based on its signature (go-onto-target) and projectiles designed to hit a particular geographic location where the target is located (go-onto-location-in-space).

Weapons designed to hit a particular target (go-onto-target) are conceptually similar to the Wren torpedo; they home in on moving targets with a particular signature. Some use passive sensors, merely sensing signals from the environment such as noise, heat or electromagnetic signals and adjusting direction as needed to close in on a target. Others have active sensors that send out signals, such as radar, and look for a target’s return. Some “lock on” to the target before launch. Others are launched at a target, like the Wren, then travel a distance and activate their seeker in order to more accurately home in on the target in the endgame, “locking on” after launch. This would be the case if the onboard seeker on the munition did not have sufficient range to detect the target at the distance it was launched. Over time, seeker processors have become increasingly sophisticated as militaries have competed in a constant race to field countermeasures to confuse and distract adversary seekers, while militaries have simultaneously improved the ability of seekers to interpret signals in the environment and distinguish countermeasures from true targets.

In contrast, go-onto-location-in-space guided munitions are not designed to hit a particular target based on its signature, but rather a particular point in geographic space where a target is known to be located. For example, GPS-guided bombs fly to a particular point based on signals from the satellite-based global positioning system. Cruise missiles frequently use terrestrial guidance systems that sense the ground beneath the missile, typically via a radar altimeter, and compare it to an onboard map, allowing the missile to fly a pre-programmed route. Laser-guided bombs fly toward a spot designated by a laser pointer, either from an aircraft or troops on the ground. In addition, submarine-launched ballistic missiles use celestial guidance, or star-matching, for navigation.

Guided munitions also vary in other important ways:

- Some guided munitions have the ability to be controlled, aborted or retargeted in flight. For example, the latest version of the Tomahawk land-attack cruise missile, the Block IV, has a communications link that allows in-flight retargeting.


27. “Fire and forget” has, of course, been the norm for projectiles in war since the sling and stone. It is only in very recent modern times with the invention of precision-guided weapons with communications links that humans have had the ability to retarget or abort a projectile while in flight.
striking a target could be significant for some types of weapons, up to several hours for cruise missiles.\textsuperscript{28} This means that human controllers may not have real-time, high-fidelity awareness of the target immediately before the moment of impact. Instead, decisions about the use of force are based on the nature of the target and a general understanding of the likelihood of potential collateral damage or civilian casualties, given the target’s location.\textsuperscript{29}

- In some cases, multiple guided weapons might be launched at a group of targets. For example, an aircraft pilot might launch multiple air-to-air guided missiles at a group of oncoming enemy aircraft.
- In other cases, some weapons have multiple guided sub-munitions. For example, a sensor fuzed weapon releases smaller sub-munitions, each of which have onboard seekers to home in on a specific group of enemy tanks.\textsuperscript{30}

As autonomy becomes more sophisticated, it may be possible for a weapon to identify a specific tank, ship or airplane because of a unique signature of that particular vehicle. In general, however, guided munitions are used to strike only the specific target or targets that a person has chosen because the freedom of maneuver of the munition is fairly constrained. Go-onto-location-in-space weapons go to a particular place chosen by a human. Go-onto-target weapons that “lock on” before launch are already locked onto a specific target chosen by the human controller before they are released. Go-onto-target weapons that “lock on” after launch are constrained in time and space in order to ensure they will engage only the target chosen by the human controller. They are not used to search over a wide area for a target or loiter until a target arrives, but rather are launched in a particular direction at a specific target. The field of view and loiter time of the seeker is relatively constrained, with the intent of ensuring that only the desired target is within the seeker’s field of view when the seeker activates.\textsuperscript{31}

Consider the example of the Wren acoustic homing torpedo. World War II submariners did not simply launch the Wren wildly into the deep blue of the ocean to go searching for ships. To do so would have been militarily ineffective, wasting a torpedo. A submarine would patrol and, once it found an enemy ship, launch the torpedo at the ship. At this stage, the Wren was no different from the types of unguided projectiles, such as rocks and arrows, that have characterized warfare for millennia. After 400 meters, the Wren would activate its acoustic homing seeker, however, which would allow the Wren to more accurately zero in on the enemy ship, striking the desired target.\textsuperscript{32} In this sense, the Wren functioned very similar to an unguided projectile in terms of how it was used; it simply had the addition of a sensor that allowed it to correct for aiming errors and home in on targets if it was off-course.

**IDENTIFYING, ACQUIRING, TRACKING, CUEING AND PRIORITIZING TARGETS**

Modern militaries also use autonomy in several other engagement-related functions prior to weapons release. Radars and other sensors help acquire, track, identify and cue potential targets to human operators.

\textsuperscript{28} See Watts, “Six Decades of Guided Munitions and Battle Networks: Progress and Prospects.” This is more likely to be the case for go-onto-location-in-space munitions since, for go-onto-target munitions, the target will likely move over a long enough time period.

\textsuperscript{29} This is not a new development. Humans have been using weapons where they do not have real-time sight of the target area since at least the invention of the catapult.


\textsuperscript{31} This is why it is particularly important to examine not only the technical specifications of the weapon, but also the doctrine and procedures for its use.

\textsuperscript{32} LLC Books, eds., Naval Weapons of Germany.
For some sophisticated weapons that require precise engagement, the timing of precisely when to fire the weapon and/or when to detonate are automated.\footnote{Watts, “Six Decades of Guided Munitions and Battle Networks: Progress and Prospects.”}

When used in conjunction with guided munitions, current uses of force can therefore involve a high degree of automation in a variety of engagement-related functions, though the decision of which specific target or group of targets are to be engaged is still made by a person.

Consider, for example, an air-to-air missile engagement. In the case of beyond visual range engagements, potential targets are identified by a computer which conveys this information to the pilot; the pilot does not have the ability to visually confirm the identity of the target, nor does the pilot even have a picture image of the target. Instead, the pilot is relying on information fed from a computer, typically based on radar. Based on this data as well as other information such as altitude, airspeed, identification friend or foe (IFF) signals and an understanding of the overall situation, the pilot makes a determination whether it is appropriate to engage this particular aircraft or group of aircraft. If the pilot decides that it is appropriate to engage, then the pilot launches an air-to-air homing missile at the enemy aircraft. If there are multiple aircraft to be engaged in the same area, the pilot might launch multiple missiles at the group of aircraft. If the missiles function with a lock-on-before-launch feature, then specific target data will be passed to the missile before launch. If the missile is functioning in a lock-on-after-launch capacity, then the missile does not have any specific targeting data on the particular enemy aircraft to be engaged. Instead, the missile will fly to a point in space and then activate its seeker, looking for targets that meet its parameters. The pilot ensures that the missile only engages the targets that he or she intends to engage by the use of tactics, techniques and procedures to ensure that, when the missile activates its seeker, the only targets within the seeker’s field of view are those that the pilot intends to engage.

Note that in this situation, the missile is “autonomous” in the sense that, once it is fired, the pilot does not have the ability to recall or control the missile. There is also a high degree of automation in numerous stages of the decision cycle. However, a person decides which specific targets to engage. The weapon is not designed to search over a wide area and select its own targets. Instead, the weapon is designed and used to home in on the individual target or specific group of targets that the pilot intends to be engaged.

**PRECISION-GUIDED WEAPONS INCREASE HUMAN CONTROL OVER WARFARE**

Since the advent of the Wren homing torpedo in World War II, guided munitions have become an increasingly important component of modern militaries. One of the earliest prominent uses of air-to-ground guided munitions by the United States military, for example, was in Operation Linebacker in Vietnam.\footnote{See Kenneth P. Werrell, Chasing the silver bullet: U.S. Air Force weapons development from Vietnam to Desert Storm (Washington, DC: Smithsonian Books, 2003).} Their use, alongside automated target identification radars and other autonomous functions, has given humans greater, not lesser, control over the use of force in warfare.\footnote{Koplow, Death by Moderation: The U.S. Military’s Quest for Useable Weapons.} This is because they have made weapons more precise, allowing militaries to target particular buildings or even floors of buildings, rather than carpet-bombing whole cities as combatants did before the advent of precision-guided munitions. In fact, some human rights groups have suggested that not using precision-guided weapons when dropping bombs in an urban area could be a war crime.\footnote{“Ukraine: Unguided Rockets Killing Civilians” (Human Rights Watch), July 24, 2014, www.hrw.org/news/2014/07/24/ukraine-unguided-rockets-killing-civilians.}
Next-generation precision-guided weapons continue this trend, leveraging automation to help humans engage specific targets that humans have decided should be engaged. While some activists have raised concerns about these weapons, more sophisticated autonomy in and of itself is not a bad thing. More advanced automated target recognition could be used to help ensure that a munition hits the target selected by the human operator, for example. Rather, the key place to focus attention is which tasks are being automated and which does the human retain. Weapon systems that incorporate autonomy to more accurately or effectively engage specific targets or clusters of targets that have been chosen by a person do not raise substantively new issues, but rather are a continuation of the way in which autonomy has been used for over seventy years.

**Human “on the loop” for selecting and engaging specific targets**

At least 30 nations use human-supervised defensive systems with greater autonomy, where humans are “on the loop” for selecting and engaging specific targets. To date, these have been used for defensive situations where the reaction time required for engagement is so short that it would be physically impossible for humans to remain “in the loop” and take positive action before each engagement and still defend effectively. Instead, autonomy is used to complete the engagement against incoming threats that meet certain criteria according to pre-programmed rules determined by humans. Human controllers are cognizant of the specific targets being engaged and can intervene to halt the weapon system, but do not make a positive decision to engage specific targets.

Currently deployed defensive, human-supervised autonomous weapon systems include:

- Air and missile defense systems, such as the U.S. ship-based Aegis and the land-based Patriot systems.
- Shorter-range counter-rocket, artillery and mortar systems to defend land bases, such as the Dutch Goalkeeper system.
- Active protection systems for ground vehicles to automatically engage incoming rockets or missiles, such as the French Shark or Russian Arena systems.

At least 30 nations employ or have in development at least one system of this type, including: Australia, Bahrain, Belgium, Canada, Chile, China, Egypt, France, Germany, Greece, India, Israel, Japan, Kuwait, the Netherlands, New Zealand, Norway, Pakistan, Poland, Portugal, Qatar, Russia, Saudi Arabia, South Africa, South Korea, Spain, Taiwan, the United Arab Emirates, the United Kingdom, and the United States.

In all of these cases, automation is used to defend human-occupied bases or vehicles from being overwhelmed by a rapid barrage of missiles or rockets. These systems are an essential component of defending and surviving in modern warfare, and automated missile defense systems can also serve an invaluable role in defending civilian populations from attack.

These systems have, however, been used very narrowly to date. They are employed to defend human-occupied vehicles and installations; they only target objects (missiles, rockets or aircraft), not people; and human

---

38. See Appendix B
39. See Appendix B.
controllers supervise the system’s operation in real-time and can intervene if necessary. Given the short time required for engagements, human controllers may not be able to intervene before any inappropriate engagements occur, but they can terminate further functioning of the system. Moreover, because human controllers are physically co-located with the system, either on a vehicle or a land base, they have physical access and can exercise hardware-level overrides in the event of a software malfunction or cyber attack.

**Human “out of the loop” for selecting and engaging specific targets**

There are a limited number of existing weapons that have a human fully out of the loop for selecting specific targets to be engaged, such that the weapon itself is selecting the targets. These weapon systems use autonomy to engage general classes of targets in a broad geographic area according to pre-programmed rules, and humans controllers are not aware of the specific targets being engaged.

**LOITERING MUNITIONS**

One type of weapon that selects and engages targets on its own, without a human “in” or “on the loop” is a loitering munition. Loitering munitions differ from guided munitions in that they are not launched at a specific target. Rather, they are launched into a general area where they will loiter, flying a search pattern looking for targets within a general class, such as enemy radars, ships or tanks. Then, upon finding a target that meets its parameters, the weapon will fly into the target and destroy it. Thus, unlike a homing munition, the human operator launching the weapon does not know the specific target that is to be engaged, only that the weapon will engage targets of a particular class within a broad geographic area.

Examples of experimental loitering munitions include the U.S. low-cost autonomous attack system (LOCAAS), which was designed to target tanks, and Tacit Rainbow, a loitering anti-radar munition. Neither was ever deployed by the United States military. The United States did have in its inventory a loitering anti-ship missile in the 1980s called the Tomahawk Anti-Ship Missile (TASM). It was removed from the U.S. inventory in the 1990s.

The only currently operational loitering munition, according to extensive research, appears to be the Israeli Harpy. The Harpy is an anti-radar weapon that flies a search pattern over a wide area searching for enemy targets.

---


radars. If it detects any radar that meets its criteria, the Harpy then dive-bombs into the radar, destroying it. 45

In a technical sense, loitering munitions chiefly differ from guided munitions in terms of their time of flight and the geographic area they cover. The sensor and algorithm used for a loitering munition may not be any more sophisticated than that of a guided munition. A lock-on-after-launch guided munition may similarly only be able to recognize a particular class of targets. The differences in terms of loiter time and geographic area result in a significant difference in how they are used, however, and in the role of the human controller in terms of selecting specific targets.

Because the freedom of a guided munition is limited in both time and space, guided munitions are only useful if launched at specific targets. Otherwise, if the guided munition activates its seeker and there is not a target directly in its field of view, the munition will likely run out of fuel and fail since it cannot search a wide area to find a target. Loitering munitions, on the other hand, are designed to search over a wide area to find a target. This is a critical distinction: for guided munitions, the human controller must know the specific target to be engaged – otherwise the munition is useless. 46 Loitering munitions are launched without a particular target, or particular target location, in mind.

Similarly, compare a loitering munition to the sensor fuzed weapons described above. While the sensor fuzed weapon deploys multiple homing sub-munitions against a group of tanks, the weapon as a whole is launched against a specific group of enemy vehicles located in a particular area. In fact, it would be difficult to use sensor fuzed weapon any other way, since the time of flight of the weapon is extremely short. If a sensor fuzed weapon were launched into a general area where there was not knowledge of specific enemy tanks then, just like other homing munitions, it would be useless because sensor fuzed weapon cannot loiter or search over a geographic area of several kilometers. For sensor fuzed weapon, therefore, even though it can attack multiple targets simultaneously, the human launching the weapon must know exactly which specific group of vehicles is to be attacked, or the weapon is useless. This highlights the importance of looking not only at the sophistication of the weapon’s seeker, but the entirety of the weapon system and its intended use.

ENCAPSULATED TORPEDO MINES
Mines are generally excluded from discussions about autonomous weapons. While there is clearly no person in the loop when the mine detonates, mines are extremely simple devices, and therefore predictable, and a person often emplaces the mine at a particular place, so the location of the mine is known. With remotely delivered mines, the precise location of each mine is not known, but the mined area is known and the mines are immobile. Thus, human control of mines is different than with other weapons, but the mine itself still has very limited freedom of action. 47 A human may not know the specific target that will detonate the mine,

45. The Harpy is frequently referred to as an unmanned aerial vehicle (UAV) or unmanned combat air vehicle (UCAV), but the Harpy is essentially a wide-area loitering cruise missile, so it should not be grouped with recoverable drones like the remotely-piloted unmanned aircraft flown by many nations.
46. For example, a Harpy with a maximum range of 500 kilometers and a top speed of 185 kilometers per hour has a time of flight of at least 2.7 hours. This allows it to loiter over a wide area to search for targets. A high-speed anti-radiation missile (HARM), on the other hand, similarly is used to destroy enemy radars but is a homing munition and must be launched at a specific enemy radar in order to be useful. With a maximum range of 90 kilometers and a top-speed of over 1,200 kilometers per hour, the HARM is only airborne for approximately four and a half minutes. Thus, it would not be practicable to use a HARM as a loitering wide-area weapon. See Robert O’Gorman and Chriss Abbott, “Remote Control War: Unmanned combat air vehicles in China, India, Israel, Iran, Russia, and Turkey” (Open Briefing, September 2013), 75; and United States Navy, “AGM-88 HARM Missile,” February 20, 2009, http://www.navy.mil/navydata/fact_display.asp?cid=2200&tid=300&ct=2.
47. Accurate record keeping and maintenance of control over the territory where the mine is located is also necessary to maintain control over the mine.
but the human controller will know the specific location.48

Encapsulated torpedo mines are a special case, however.49 Encapsulated torpedo mines are a type of sea mine that, when activated by a passing ship, instead of exploding, open a capsule that then releases a torpedo that engages a target. In this case, the torpedo is not being used to home in on a target that has been selected by a person. Nor is the mine simply blowing up in place. Instead, the mine is given a greater freedom of maneuver to release a torpedo that will then track onto a nearby target. De facto, the mine is selecting and engaging targets on its own. Thus, encapsulated torpedo mines are, arguably, more similar to loitering munitions than to traditional mines.

Encapsulated torpedo mines were part of the U.S. inventory for a number of years, but have since been retired. Russia and China have an encapsulated torpedo mine, the PMK-2, in service today.50

IV. WHAT IS AN “AUTONOMOUS WEAPON?”

As previously mentioned, there is no internationally agreed-upon definition for an “autonomous weapon,” which complicates discussion of the topic. From a common sense perspective, it makes sense to define an “autonomous weapon” as one that selects and engages targets on its own, and a “semi-autonomous weapon” as one where a human is selecting the specific targets for engagement. These are the definitions most commonly used in the literature on autonomous weapons and are appealing for a number of reasons.

First, the common sense perspective conforms to the widely held assumption that “autonomous weapons” (whatever they are defined to be) do not generally exist today, with some isolated exceptions.51 Second, it seems to capture the essence about what is so chilling about the vision of a futuristic hunter-killer robot deciding who to kill on its own, without human authorization. Third, it essentially mirrors the definitions already used by a diverse set of organizations that have weighed in on autonomous weapons, including the U.S. Department of Defense, Human Rights Watch, the UN Special Rapporteur on extrajudicial, summary or arbitrary executions, and the International Committee of the Red Cross (ICRC).52

Common sense definitions may not be sufficient, however, due to interest in setting exactly where the line is between autonomous and semi-autonomous weapon systems. This type of line is critical for all stakeholders in discussions of autonomy – for military planners attempting to figure out how to direct research and development programs, as well as for activists thinking about what types of activities they wish to place under greater scrutiny.

Additionally, the most likely near-term candidates for autonomous weapons are not malevolent, sentient

48. From a certain perspective, mines could be thought of as a loitering go-onto-location-in-space munition.
51. The defensive, human-supervised autonomous weapon systems in use by at least 30 nations today are a notable exception, although they are used in a very restricted context, with real-time human supervision, and physical access to disable the system if necessary. Human Rights Watch has referred to these defensive systems as “precursors” of the types of weapons they fear, which this definition seems to accurately capture. The Harpy loitering munition and PMK-2 encapsulated torpedo mine are isolated exceptions of fully autonomous weapon systems under this definition.
52. See Appendix A.
humanoid robots (like those of movie fame), but rather something more like wide-area search-and-destroy loitering munitions. Thus, the definitions must clearly distinguish, in a way that is technically rigorous, between autonomous weapons and the precision-guided homing munitions that have been in use for over seventy years.\(^{53}\)

From a more technical perspective, the following definitions offer a clearer way forward:

- An **autonomous weapon system** is a weapon system that, once activated, is intended to select and engage targets where a human has not decided those specific targets are to be engaged.

- A **human-supervised autonomous weapon system** is a weapon system with the characteristics of an autonomous weapon system, but with the ability for human operators to monitor the weapon system's performance and intervene to halt its operation, if necessary.

- A **semi-autonomous weapon** is a weapon system that incorporates autonomy into one or more targeting functions and, once activated, is intended to only engage individual targets or specific groups of target that a human has decided are to be engaged.

These definitions are intended to clarify and sharpen the distinction between autonomous and semi-autonomous weapons in two important ways:

First, the definition of an autonomous weapon system is intended to clearly delineate that autonomous weapon systems select and engage targets where a human has not decided those specific targets are to be engaged. This essentially spells out what is meant by saying an autonomous weapon would select and engage targets “on its own.” At the risk of redundancy, this language’s focus on the role of the human operator is designed to clarify any confusion surrounding guided munitions, particularly those that use seekers to home in on targets or specific groups of targets that a human has decided are to be engaged.

Second, the idea of a human decision is embedded within each of the above definitions. This highlights what is different about an autonomous weapon from the perspective of a human operator. The decision to place an autonomous weapon into operation versus a semi-autonomous one is a very different decision. Even in the case of a fire-and-forget homing missile, while the person launching the missile may not be able to recall it, the human operator has made a decision about which individual target or specific group of targets is to be engaged by that homing missile. In the case of an autonomous weapon, however, the human has decided to launch a weapon to seek out and destroy a general class of targets over a wide area, but is not making a decision about which specific targets are to be engaged. Both definitions, however, focus on the decision the human is making or not making and do not apply the word “decision” to something the weapon itself is doing, which could raise murky issues of machine intelligence and free will.

53. We think the distinction is clear in the commonly used “select and engage” definition, but some have seized on the notion that a homing munition has a seeker to suggest that it is “selecting” targets on its own. If one were to define autonomous weapons this way, then the logical conclusion would be that autonomous weapons have been around for nearly three-quarters of a century, so this entire discussion is a lot of fuss for nothing. We think that there is a significant difference between a homing munition that is used to engage a target that has been chosen by a person and a wide-area search-and-destroy weapon that would select and engage targets on its own, so this definition is modified to further clarify that distinction. For more on this topic, see Mark Gubrud, “Killer Robots and Laser-Guided Bombs: A reply to Horowitz & Scharre”; John Markoff, “Fearing Bombs That Can Pick Whom to Kill”; Rebecca Crooteof, “The Killer Robots Are Here: Legal and Policy Implications,” Cardozo Law Review, 36 (2015), papers.ssrn.com/sol3/papers.cfm?abstract_id=2534567; Michael C. Horowitz and Paul Scharre, “Do Killer Robots Save Lives?” Politico Magazine, November 19, 2014, www.politico.com/magazine/story/2014/11/killer-robots-save-lives-113010.html#.WHODj1XOKY8; “Missile Systems and Human Control” (Campaign to Stop Killer Robots).
It should be noted that because of the different ways that the word autonomy is used, there are many ways a semi-autonomous weapon system could become “more autonomous” in a general sense and still not be an “autonomous weapon system.” A semi-autonomous weapon could incorporate greater autonomy into other functions, such as being able to avoid threats or change its route, without changing human control over the decision for which specific targets are to be engaged. Also, a weapon system could become “smarter,” incorporating more capable sensors and seekers, but still retain human control over the selection of specific targets. Thus, a semi-autonomous weapon that had automatic target recognition technology and used that technology to ensure that it only engaged individual targets or specific groups of targets that a person had decided are to be engaged would still be a semi-autonomous weapon. This reinforces the importance of looking not only at the weapon’s seeker but the entire capability of the weapon system and its intended use.

An alternative terminology for understanding how autonomy is used in weapons – and which functions are most important – could be to focus on the function being automated. In this case, autonomous weapons would be “self-targeted weapons,” indicating that the weapon itself is the one choosing the specific target to be engaged. Semi-autonomous weapons would be “human-targeted weapons,” making clear that, while other functions may use autonomy, the human is still choosing the weapon. This alternative terminology may be clearer and could help avoid some of the confusion stemming from the word “autonomy.”

V. CONCLUSION

Rapid advances in information technology means that capabilities that may seem at present to be distant or impossible may come to fruition sooner than we realize. Activist groups that have catalyzed this debate are to be commended for raising awareness on an important issue. However, leaping to the conclusion that a legally-binding ban on autonomous weapons is required before better understanding what autonomous weapon systems are and the issues they raise is putting the cart before the horse. Better understanding of the issue is needed before deciding on a path forward.

This paper is intended to frame discussions about autonomy in weapons in order to help inform ongoing debates about the appropriate role of autonomy and human control in the use of force. Too often, arguments about how autonomy ought to be used in future weapons are made without a good understanding of how autonomy is already used today and, in some cases, has been used for decades. In some situations, this can mean individuals advocating for rules governing autonomy in future weapons that would, if implemented, ban wide swaths of existing weapons without any clear evidence that such weapons have been harmful. In fact, in some forms of autonomy, such as those that enable precision-guided weapons, have actually reduced civilian casualties in war. In any case, rules that capture weapon systems that have already been used widely for decades without significant concerns almost certainly miss the mark about what is novel about potential future autonomy in weapons.

Future weapon systems are likely to incorporate greater autonomy in a number of ways, and different types of

54. For example, the International Committee for Robot Arms Control’s (ICRAC) proposed “minimum necessary standards” for meaningful human control would, if implemented strictly, ban huge classes of weapons that have been used for centuries dating back to the invention of the catapult. International Committee for Robot Arms Control, “ICRAC statement on technical issues to the 2014 UN CCW Expert Meeting, May 14, 2014, http://icrac.net/2014/05/icrac-statement-on-technical-issues-to-the-un-ccw-expert-meeting/.

autonomy raise different concerns. Some types of autonomy, such as the ability to maneuver around obstacles or threats, or to deploy evasive maneuvers or non-lethal countermeasures like flares or chaff, raise few if any significant concerns. Others raise potentially profound concerns about the role of humans in the use of force.

Thus, clear terminology that delineates between existing weapon systems that incorporate some autonomy but retain human control over the specific targets being engaged and future autonomous weapon systems that would select and engage targets on their own is an essential precondition for debate. Without a commonly agreed-upon lexicon, those discussing the topic of autonomous weapons may fall victim to the different ways in which the word “autonomy” is used.

It also is worth noting that while this paper has focused on autonomy in engagement decisions for weapon systems, there are many other features of the system that may be equally if not more important. These could be related to autonomy in other, non-engagement related aspects of the system, or not related to autonomy at all.

For example, a static autonomous weapon that could select and engage targets on its own (stationed along a country’s border, for instance), would have a significantly different level of risk than a mobile system. And a mobile autonomous weapon system that can only operate in a limited geographic area would have a different level of risk than one that had the ability to patrol over a wide area. Similarly, non-autonomous features could also affect the costs and benefits of utilizing an autonomous weapon. Some environments, such as undersea, are less cluttered and have less potential for civilian casualties. Another important consideration is the target being engaged. Autonomous weapon systems that target objects (i.e. anti-materiel or anti-vehicle autonomous weapons) versus people (i.e. anti-personnel autonomous weapons) also pose meaningfully different levels of risk from the perspective of distinguishing military from civilian targets and avoiding civilian casualties.

Finally, the destructive capability of the weapon itself is a significant factor in thinking about risk. A bomber has the capability to cause significantly more destruction in the event of a malfunction or cyber attack than a machine gun. All of these factors are significant and come to bear on how to think about the relative potential costs and benefits of autonomous weapon systems.

How autonomy has been used to date highlights the importance of some of these distinctions. While at least thirty nations operate human-supervised autonomous weapons, they are used in a relatively limited context. They have been used, to date, for static defense of installations or onboard defense of human-occupied vehicles. They target objects (anti-vehicle/materiel), not people (anti-personnel). Perhaps most importantly, human controllers not only supervise their operation in real time but also have physical access to the system so that they can manually disable the system in the event of a malfunction, communications failure or cyber attack.

These issues raise important questions for future research. This paper informs current discussions about autonomous weapons by clarifying terminology and helping focus attention on how potential new uses of autonomy in future weapons could differ from historical uses of autonomy. This should be a launching point for further much-needed discussions among activists, lawyers, ethicists, academics, engineers, and military and defense professionals on the appropriate role of autonomy in future weapon systems.
Appendix A: Definitions of Autonomous Weapons from Selected Reports

The U.S. Department of Defense (DoD) official policy on autonomy in weapon systems, DoD Directive 3000.09, *Autonomy in Weapon Systems*, uses the following definitions:56

**autonomous weapon system**: A weapon system that, once activated, can select and engage targets without further intervention by a human operator.

This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.

**human-supervised autonomous weapon system**: An autonomous weapon system that is designed to provide human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur.

**semi-autonomous weapon system**: A weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator.

This includes:

Semi-autonomous weapon systems that employ autonomy for engagement-related functions including, but not limited to, acquiring, tracking, and identifying potential targets; cueing potential targets to human operators; prioritizing selected targets; timing of when to fire; or providing terminal guidance to home in on selected targets, provided that human control is retained over the decision to select individual targets and specific target groups for engagement.

“Fire and forget” or lock-on-after-launch homing munitions that rely on tactics, techniques, and procedures to maximize the probability that the only targets within the seeker’s acquisition basket when the seeker activates are those individual targets or specific target groups that have been selected by a human operator.

**target selection**: The determination that an individual target or a specific group of targets is to be engaged.

Christof Heyns, UN Special Rapporteur on extrajudicial, summary or arbitrary executions, wrote in his April 2013 report:57

*The measure of autonomy that processors give to robots should be seen as a continuum with significant human involvement on one side, as with UCAVs where there is “a human in the loop”, and full autonomy on the other, as with [lethal autonomous robotics (LARs)] where human beings are “out of the loop.*

Under the currently envisaged scenario, humans will at least remain part of what may be called the “wider loop”: they will programme the ultimate goals into the robotic systems and decide to activate and,

---

if necessary, deactivate them, while autonomous weapons will translate those goals into tasks and execute them without requiring further human intervention.

Supervised autonomy means that there is a “human on the loop” (as opposed to “in” or “out”), who monitors and can override the robot’s decisions. However, the power to override may in reality be limited because the decision-making processes of robots are often measured in nanoseconds and the informational basis of those decisions may not be practically accessible to the supervisor. In such circumstances humans are de facto out of the loop and the machines thus effectively constitute LARs.

“Autonomous” needs to be distinguished from “automatic” or “automated.” Automatic systems, such as household appliances, operate within a structured and predictable environment. Autonomous systems can function in an open environment, under unstructured and dynamic circumstances. As such their actions (like those of humans) may ultimately be unpredictable, especially in situations as chaotic as armed conflict, and even more so when they interact with other autonomous systems.

The terms “autonomy” or “autonomous”, as used in the context of robots, can be misleading. They do not mean anything akin to “free will” or “moral agency” as used to describe human decision-making. Moreover, while the relevant technology is developing at an exponential rate, and full autonomy is bound to mean less human involvement in 10 years’ time compared to today, sentient robots or strong artificial intelligence are not currently in the picture.

Human Rights Watch, in their 2012 report “Losing Humanity,” used the following terminology:

Although experts debate the precise definition, robots are essentially machines that have the power to sense and act based on how they are programmed. They all possess some degree of autonomy, which means the ability of a machine to operate without human supervision. The exact level of autonomy can vary greatly. Robotic weapons, which are unmanned, are often divided into three categories based on the amount of human involvement in their actions:

- **Human-in-the-Loop Weapons**: Robots that can select targets and deliver force only with a human command;
- **Human-on-the-Loop Weapons**: Robots that can select targets and deliver force under the oversight of a human operator who can override the robots’ actions; and
- **Human-out-of-the-Loop Weapons**: Robots that are capable of selecting targets and delivering force without any human input or interaction.

The International Committee of the Red Cross (ICRC), in a report on an expert meeting held in March 2014 on autonomous weapons, explained:

There is no internationally agreed definition of autonomous weapon systems. For the purposes of the meeting, ‘autonomous weapon systems’ were defined as weapons that can independently select and

---

58 “Losing Humanity: The Case Against Killer Robots” (Human Rights Watch).
Appendix B: Selected Examples of Human-Supervised Autonomous Weapon Systems

SELECTED EXAMPLES OF AIR AND MISSILE DEFENSE SYSTEMS

AEGIS COMBAT SYSTEM
Date of Introduction ................. 1983¹
Manufacturer ......................... Lockheed Martin²
Manufacturing Country .......... United States
National Operators ................. Australia, Japan, Norway, South Korea, Spain, and the United States³


GOALKEEPER
Date of Introduction ................. Early 1990s¹
Manufacturer ......................... Thales
Manufacturing Country .......... Netherlands²
National Operators ................. Belgium, Chile, the Netherlands, Portugal, Qatar, South Korea, the United Arab Emirates, and the United Kingdom³

3. Ibid.

KASHTAN CLOSE-IN WEAPON SYSTEM
Date of Introduction ................. 1988¹
Manufacturer ......................... KBP Instrument Design Bureau²
Manufacturing Country .......... Russia
National Operators ................. China, India, Russia³


MK 15 PHALANX CLOSE-IN WEAPON SYSTEM
Date of Introduction ................. 1980¹
Manufacturer ......................... The Raytheon Company²
Manufacturing Country .......... United States
National Operators ................. 25 navies including Australia, Bahrain, Canada, Egypt, India, Israel, Japan, New Zealand, Pakistan, Poland, Portugal, Saudi Arabia, South Korea, Taiwan, the United Kingdom, and the United States³

2. The Raytheon Company, “Phalanx Close-In Weapon System (CIWS).”

Patriot
Date of Introduction ................. 2002¹
Manufacturer ......................... The Raytheon Company²
Manufacturing Country .......... United States
National Operators ................. Egypt, Germany, Greece, Israel, Japan, Kuwait, the Netherlands, Saudi Arabia, South Korea, Spain, Taiwan, the United Arab Emirates, and the United States³
2. The Raytheon Company, “Patriot.”

SeaRAM ANTI-SHIP MISSILE DEFENSE SYSTEM
Date of Introduction .................. 2008\(^1\)

Manufacturer ......................... The Raytheon Company\(^2\)

Manufacturing Country .......... United States

National Operators ............... Japan and the United States\(^3\)


SELECTED EXAMPLES OF GROUND VEHICLE ACTIVE PROTECTION SYSTEMS

AMAP-ADS
Date of Introduction .................. 2011\(^1\)

Manufacturer ......................... IBD Deisenroth Engineering\(^2\)

Manufacturing Country .......... Germany

National Operators ............... “An Asian nation” (rumored to be Singapore)\(^3\)

1. Rheinmetall Defence, “Rheinmetall takes up a majority share in ADS GmbH,” January 2, 2011.
2. IBD Deisenroth Engineering, “AMAPTM-ADS: The Active Protection System.”

Date of Introduction ............... 1993\(^3\)

Manufacturer ...................... KBP Instrument Design Bureau\(^2\)

Manufacturing Country .......... Russia

National Operators .............. Russia\(^3\)

2. “ARENA Active Protection System.”

DROZD
Date of Introduction .................. 1983\(^1\)

Manufacturer ...................... KBP Instrument Design Bureau\(^2\)

Manufacturing Country .......... Russia

National Operators .............. Russia\(^3\)

2. Ibid.
3. Ibid.

DROZD-2
Date of Introduction .................. 2005\(^1\)

Manufacturer ...................... KBP Instrument Design Bureau\(^2\)

Manufacturing Country .......... Russia

National Operators .............. Russia\(^3\)

2. Ibid.
3. Ibid.

IRON CURTAIN
Date of Introduction ............... In development\(^1\)

Manufacturer ...................... Artis, LLC\(^2\)

Manufacturing Country .......... United States
National Operators .................. United States

3. Hambling, “New Bomb-Resistant Trucks Will Blast RPGs Before They Hit.”

LEDS-150
Date of Introduction ............... 2015 (anticipated)\(^1\)

Manufacturer ........................ Saab\(^2\)

Manufacturing Country .......... South Africa

National Operators ................. Pakistan and India are reportedly interested in purchasing the LEDS-150\(^3\)

2. aab, “LEDS full spectrum active protection for land vehicles.”

QUICK KILL
Date of Introduction ............... In development\(^1\)

Manufacturer ........................ The Raytheon Company\(^2\)

Manufacturing Country .......... United States

National Operators .................. United States\(^3\)

1. The Raytheon Company, “Raytheon’s Quick Kill Active Protection System defeats one of the most lethal armor-piercing Rocket Propelled Grenades,” January 9, 2013.
2. The Raytheon Company, “Quick Kill” Active Protection System (APS).”
3. The Raytheon Company, “Raytheon’s Quick Kill Active Protection System defeats one of the most lethal armor-piercing Rocket Propelled Grenades.”

SHARK
Date of Introduction ............... 2006 (est.)\(^3\)

Manufacturer ........................ Thales Group and IBD Deisenroth Engineering\(^2\)

Manufacturing Country .......... France\(^3\)

National Operators .................. France\(^4\)

3. Ibid.
4. Ibid.

TRENCH COAT
Date of Introduction ............... 2012 (est.)\(^1\)

Manufacturer ........................ Rafael Advanced Defensive Systems\(^2\)

Manufacturing Country .......... Israel

National Operators .................. Israel\(^3\)

2. Ibid.

TROPHY
Other Name .......................... ASPRO-A

Date of Introduction ............... 2007\(^1\)

Manufacturer ........................ Rafael Advanced Defensive Systems\(^2\)

Manufacturing Country .......... Israel

National Operators .................. Israel\(^3\)

2. Rafael Advanced Defensive Systems, “Trophy.”
Great Power Competition and the AI Revolution: A Range of Risks to Military and Strategic Stability

By Elsa Kania

Today’s rapid advances in artificial intelligence (AI) could disrupt and destabilize the existing military balance—and not necessarily for the reasons that have captured popular imagination. The potential realization of Artificial General Intelligence or “superintelligence” merits discussion, but it remains a relatively distant possibility. Yet advances in AI have already started to provoke fear, and often hyperbole, about the threats of “killer robots,” a looming Terminator Dilemma and the risk of an AI arms race. Earlier this month Elon Musk even tweeted, “Competition for AI superiority at national level most likely cause of WW3 [in my opinion].”

These fears may be premature, but AI’s disruptive potential is real. The recent progress in AI has primarily involved machine learning and particularly deep learning techniques, such as the use of deep neural networks, across disciplines including computer vision, pattern recognition and natural language processing. Since 2016, several critical milestones have revealed the rapid pace of advances and potential real-world applications. In 2016, the victory of “Mayhem” in DARPA’s Cyber Grand Challenge demonstrated the potential of autonomous detection and patching of software vulnerabilities to transform cybersecurity. With the computer program AlphaGo’s historic defeat of Lee Sedol in their 2016 match and subsequent victory over Ke Jie, the world’s top Go player, AI has achieved mastery of the game of Go, which requires complex strategizing, at least a decade earlier than expected.

Recognizing the disruptive, even revolutionary implications of AI for national defense, the United States, China and Russia are actively seeking to advance their capabilities to employ AI for a range of military applications. In spring 2017, the Department of Defense (DoD) revealed it had established an Algorithmic Warfare Cross-Functional Team “to accelerate DoD’s integration of big data and machine learning.” This summer, China released the New Generation AI Development Plan, which articulated the ambition to “lead the world” in AI by 2030. This plan calls for military-civil fusion in AI to leverage dual-use advances for applications in national defense, including in support of command decision-making, military deduction, and defense equipment. Meanwhile, the Russian military has been aggressively advancing its efforts in intelligent robotics, and Russian President Vladimir Putin recently declared, “Whoever becomes the leader in [AI] will become the ruler of the world.” Indeed, the advent of AI in warfare appears to be resulting in a transformation of the character of conflict beyond information-age warfare toward “algorithmic warfare,” in the U.S. military’s phrasing, or “intelligentized” (智能化) warfare, as Chinese military thinkers characterize it.

Despite recurrent calls to ban “killer robots”—and a recent open letter that articulated concerns that the development of lethal autonomous weapons would open a “Pandora’s box” and risk unleashing “weapons of terror”—an outright ban is unlikely to be feasible. It is improbable that major powers would accept constraints on capabilities considered critical to their future military power. Even attempts to pursue some form of regulation or an international treaty to restrain military applications of AI could be readily overtaken by technological developments. The diffusion of this dual-use technology will also be difficult to control.

Consequently, major militaries should take a proactive approach to evaluating and mitigating the potential risks introduced by advances in military applications of AI. This is in their interest, as the U.S., China and Russia still share at least a basic commitment to strategic stability and recognize the undesirability of inadvertent escalation.

To date, much of the analysis and attention associated with the risks of AI in warfare has concentrated on ethical concerns and operational risks associated with the potential use of lethal autonomous weapons systems (LAWS). There has been an international debate on the challenges of applying and adapting the law of armed conflict—with core principles of necessity, distinction, proportionality and humanity—to the use of autonomous weapons systems, but it remains unclear how different militaries will interpret and abide by this traditional framework. Concurrently, as Paul Scharre of the Center for a New American Security discussed in a report on the topic, the employment of autonomous weapons could result in a range of operational risks, including inevitable failure in complex systems, adversary attempts to attack or otherwise undermine autonomous systems (e.g., spoofing and behavioral hacking), or unanticipated interactions between adversarial systems.

These issues are critical and should also inform thinking on the risks that arise out of military applications of AI beyond the context of autonomous weapons. Indeed, risks that might result from current, comparatively uncontroversial military uses of AI also deserve sustained attention. Despite fears that AI is “summoning the demon,” the present limitations of AI, rather than its potential power or questions of controllability, could become problematic in the near term. In some cases, errors and issues might arise from seemingly routine applications, even when the algorithm in question is not used directly in a weapons system or to make a life-and-death decision.
Consider, for example, the implications of rote errors for the U.S. and Chinese militaries’ progression toward using AI for the automation of intelligence, surveillance and reconnaissance (ISR) functions, particularly in support of command decision-making. With Project Maven, the Defense Department plans to accelerate its integration of big data and machine learning, using computer vision algorithms to enable the automated processing of data, video and imagery. Through this initiative, the department intends to have the capacity to “deploy” algorithms in a war zone by the end of 2017. In addition, the U.S. Air Force seeks to leverage AI to process information at speed and scale, through algorithms and human–machine interfaces, and to enhance leaders’ decision-making capabilities. Similarly, the Chinese People’s Liberation Army (PLA) is developing algorithms that enable data fusion, enhance intelligence analysis, and support command decision-making. For instance, the PLA is funding research and development related to target recognition, as well as the processing of sensor data and satellite imagery, based on machine learning. In particular, the Chinese army is focused on the potential of AI in operational command and assistance to decision-making.

These military applications of AI do not directly involve decisions about lethal force but could cause errors that might contribute to crisis instability or exacerbate the risks of escalation. At this stage of development, AI remains far from intelligent, tending to make mistakes no human would make. Such errors can be unpredictable or difficult to mitigate. In certain cases, the results can be amusing or nonsensical. In a military context, however, there could be adverse consequences, with higher likelihood of errors or unexpected emergent behavior as the level of complexity increases and if a situation exceeds the expected parameters of an algorithm.

The use of machine learning in support of military intelligence, surveillance and reconnaissance capabilities or command systems could create new avenues for misperception or miscalculation. For instance, if a computer vision algorithm used to process satellite imagery makes an error in identifying a perceived threat or potential target, the result could be a mistake in analysis or even on the battlefield. Similarly, if an algorithm used for machine translation or natural language processing incorrectly renders a critical piece of intelligence, inaccurate information could be introduced into analytical and decision-making processes. (Of course, in some cases, the use of AI could serve to mitigate suboptimal analysis and decision-making that might otherwise arise from human cognitive bias.) There is also the threat that adversaries will develop countermeasures that deliberately damage or interfere with each other’s AI systems, seeking to distort data or target the algorithm itself.

Looking to another near-term application, there is likely to be a trend toward the use of AI-enabled cyber capabilities by multiple militaries (and even non-state actors), introducing greater degrees of autonomy to a complex and contentious domain. The 2012 Defense Department directive on “Autonomy in Weapons Systems” explicitly indicated that the policy, which called for the exercise of “appropriate levels of human judgment” over use of force, did “not apply to autonomous or semi-autonomous cyberspace systems for cyberspace operations.” The sheer speed required in certain cyber operations, as in air and missile defense, could necessitate greater degrees of autonomy. In August, the Defense Department officially purchased the technology associated with Mayhem, which will be used to automate the detection and patching of software vulnerabilities. Although Mayhem is intended for a defensive mission, similar techniques could be weaponized and leveraged for offensive purposes. The Chinese military is also likely to look to AI to enhance its offensive and defensive cyber capabilities, including under the aegis of its new Strategic Support Force. The trend towards integrating AI with cyber capabilities to achieve an advantage could exacerbate escalatory dynamics in this contested domain, particularly if such capabilities start to proliferate.

Looking forward, AI will enable disruptive military capabilities while creating systemic risks to military and strategic stability. Now is the time to start considering the potential ramifications of these technological trends and evaluating appropriate parameters that might mitigate risks of error or escalation. To start, great-power militaries might consider pursuing—and perhaps, as a future confidence-building measure, committing to—an initial set of pragmatic measures, such as:

- Engaging in robust testing of the safety and integrity of military AI systems, focusing on potential errors or failures that might occur in open, uncontrolled environments;
- Creating redundancies in military AI systems, including those used to support intelligence, surveillance and reconnaissance capabilities, to enable verification, such that there are multiple methods to detect errors and evaluate outputs, in order to ensure consistency with actual ground truth;
- Exploring options for fail-safe measures or “circuit breakers” to allow for de-escalation or de-confliction in the case of unintended engagements or escalation; and
- Ensuring the “explainability” of AI systems to lessen issues of compromised decision-making, while perhaps ensuring informed, “meaningful human control” whenever feasible.
Elsa Kania is an analyst focused on the PLA's strategic thinking on and advances in emerging technologies, including unmanned systems, artificial intelligence, and quantum technologies. Elsa is also in the process of co-founding a start-up research venture. Her professional experience has included working at the Department of Defense, the Long Term Strategy Group, FireEye, Inc., Harvard's Belfer Center, and the Carnegie-Tsinghua Center for Global Policy. She is fluent in Mandarin Chinese.
PANEL III:

NATIONAL SECURITY LAW AND POLICY, FROM THE PRIVACY AND CIVIL LIBERTIES OVERSIGHT BOARD TO THE DEPARTMENT OF JUSTICE

DISCUSSANTS:
THE HONORABLE RACHEL BRAND
RAJESH DE
CRYPTO WARS 2.0: WHY LISTENING TO APPLE ON ENCRYPTION WILL MAKE AMERICA MORE SECURE

I. INTRODUCTION

Encryption is a topic that has garnered significant attention in the United States (U.S.) because of the controversy between the Federal Bureau of Investigation (FBI) and Apple. A federal magistrate judge initially ordered Apple to comply with the FBI's request that Apple break into the encrypted phone of a perpetrator of the San Bernardino terrorist attack on December 2, 2015. Apple resisted this order, going so far as to release a public letter explaining its reasons why. As per a Government Status Report filed on March 28, 2016, the FBI successfully accessed the phone in question without the assistance of Apple.

There is much debate over the role that encryption plays in the war on terror writ large, with the increased use of the Telegram application by the terrorist group known as the “Islamic State of Iraq and the Levant” (ISIS). Many Internet companies have recently moved to end-to-end encryption for messaging systems, which provides greater user privacy. In addition, many smartphone manufacturers are creating operating systems for their devices that render information stored on the devices encrypted by default. Digital encryption is a way of coding plaintext so that only the receiver of a message or information can “break the code.” But do more secure systems make us more vulnerable or do they make us safer? Furthermore, who gets to decide the norms of encryption? When Apple and Microsoft and Google are operating all over the globe, does the U.S. get to decide how communications are passed across these global networks?

Public confusion over the use of encryption technology and heavy-handed regulation by foreign states will make our communications systems more vulnerable, and by extension, less safe. The tech industry takes the position that it is not feasible to provide law enforcement agencies with a golden key to access encrypted information. They argue that compromising encryption even in the slightest would leave an opening for bad actors, whether they are hostile states, domestic hackers, or terrorist networks. U.S. policymakers, however, argue that there must be ways to access encrypted information to combat domestic and international terrorism. After the San Bernadino attacks, Apple v. FBI served as a watershed moment in the regulation of encrypted information in the U.S. and the understanding of what is at stake by the imposition of “backdoors.”
This comment begins with a brief synopsis on the technology involved in end-to-end encryption and its current uses. The comment then moves to a survey of the four main viewpoints on the utility of end-to-end encryption. This comment discusses the positions taken by those in the tech industry and argues that the majority position of those in the tech industry—that strong encryption makes systems less susceptible to hacking and cyber-attack—is the best position. Making our current encrypted systems vulnerable to any degree will open the door for hackers and hostile nations to exploit the vulnerabilities. Deferring to the positions of the tech industry while other countries make their systems more vulnerable will ultimately help make the U.S. safer from terrorism.

II. BACKGROUND ON ENCRYPTION

End-to-end encryption is a method of coding information that is sent from one user to another user, or a group of users. There are two types of basic encryption methods: public key and private key. Keys are devices or algorithms used to encode or decode messages.

Public key encryption involves a public key and a private key. A sender encodes a message using his private key that would then be decoded by the recipient's public key, or vice versa. On the other hand, private key encryption involves only one key, the private key. The sender sends the key along with the message and it is used to decode the message. Private key encryption is the less secure of the two methods of encryption because the key travels with the message.

The majority of usage of end-to-end encryption technology is for benign purposes such as to facilitate online banking and shopping. Technology (tech) companies have moved to more secure encryption technology to bolster individual privacy, and arguably to bolster their bottom line.

There is growing debate over the use of encryption technology by terrorist organizations. For example, some argue that terrorist groups used encryption technology to plan two recent terrorist attacks in Paris, France and San Bernadino, California. On November 13, 2015, in Paris, France, three suicide bombers struck outside the Stade de France, while simultaneous mass shootings were taking place at cafés, restaurants, and inside the Bataclan theatre, during which 368 people were injured and 130 people died. Officials speculated that encrypted applications (apps) such as Telegram and WhatsApp were used to plan these attacks.

On December 2, 2015, in San Bernardino, California, two individuals killed fourteen people and seriously injured twenty-two more in a terrorist attack targeted at an office holiday party. An encrypted iPhone owned by one of the perpetrators became the focal point for the controversy between the FBI and Apple over encryption and the authority of the federal government to force private companies to comply with U.S. security directives.

The increased use of encryption for both individual privacy and potential terrorist use raises the question of whether or not the increased use of more secure encryption systems makes the world safer or more vulnerable to terrorist activities and other attacks. Several major tech companies have encryption already built into their messaging systems. Many email systems or email chat systems use a form of encryption. The most common and respected form of encryption is “Pretty Good Privacy” (PGP). The founder of PGP, Phil Zimmermann, also called the “King of
Encryption,” created PGP in the early 1990s. Mr. Zimmermann initially distributed PGP for free to any Internet user that desired to use it.

True end-to-end encryption hides the content of a transmission of information from even the administrators of the application used (e.g., Apple, Facebook, etc.). End-to-end encryption is growing in popularity within the tech field due to its high degree of privacy. For example, Google is seeking to release “End-to-End,” a Chrome plug-in that encrypts messages end-to-end. Facebook has also made its own “Dark Web” website, acquired WhatsApp, a messaging company that uses end-to-end encryption in its Android application, and launched features that enabled users to get notifications through encrypted emails. Facebook appears to be moving in the direction of more encryption in its capabilities.

III. THE MAIN VIEWPOINTS ON THE DEBATE OVER ENCRYPTION

There are four main viewpoints surrounding the current debate over encryption. The first is that secure networks actually make us safer. The second is that completely secure networks force law enforcement and anti-terror groups to operate at an unacceptable handicap. The third viewpoint, which is somewhere in between, is that secure systems generally make us safer, but that exceptional access should be available in certain situations. Fourth, according to David Chaum, encryption in its purest form is not the optimal way of ensuring privacy and security, but rather encryption should be a part of a larger privacy scheme.

The tech industry generally holds the first viewpoint that completely secure networks make us safer. Those in the tech commentary community also widely believe in completely secure networks. Additionally, the Netherlands has stated this to be their national policy on regulating encryption. This viewpoint is rooted in the belief that allowing national actors to have exceptional access would create vulnerabilities that bad actors would exploit. Adherents to this viewpoint generally oppose government access on principle or out of practicality.

Law enforcement agencies around the world hold the second viewpoint that embracing full encryption presents an unacceptable risk. Tech and computer security commentators, as well as governments of certain nations, also adhere to the view that some access to encrypted information is needed. Advocates of this viewpoint emphasize findings regarding the use of encrypted devices by terrorist groups and dangerous criminals to support their view. They argue that allowing full encryption allows these bad actors to have an advantage over law enforcement. This viewpoint recognizes the benefit of government access to these communications, and that increased access is worth the accompanying privacy risks associated with less secure and more vulnerable systems.

The American national security apparatus holds the third viewpoint, which states that secure systems generally make us safer, but access in certain circumstances is necessary. National Security Agency (NSA) Director Michael S. Rogers best summarized this view: “I don't want a back door ... I want a front door. And I want the front door to have multiple locks. Big locks.” Adherents to this viewpoint argue that only certain government agencies in certain nation states should have exceptional access to encrypted communications.
The fourth viewpoint, advocated by encryption pioneer David Chaum, asserts that encryption can be part of a larger privacy scheme, but that it should not be the final arbiter of what is private and what is secure. Mr. Chaum has been creating “PrivaTegrity,” an encryption system that allows for “fully secret, anonymous communications” with a special backdoor controlled by a “council.” Chaum's system is not a pure end-to-end encryption but is a unique solution to a stalemate over privacy.

IV. THE DOMINANT POSITION OF THE TECH INDUSTRY

The majority, if not nearly all, of the big tech companies believe that end-to-end encryption is necessary to protect user data. Tech companies also believe that having encrypted systems with a “special key” or access for government and law enforcement is not technically feasible because it is costly to protect. According to the industry and surrounding commentaries, the assumption is that seeking to carve out special access for law enforcement is based on a lack of understanding of the technology involved in encryption.

There are three likely explanations to the tech industry's position on limiting access to governments and law enforcement on encryption technology. First, leaders of the large tech companies genuinely value individual privacy. Second, large tech companies genuinely think that the Internet will be the most secure--and thus our society the safest--with full encryption. Third, large tech companies believe that privacy is a good sales strategy that will increase their bottom line.

On the basis of preserving individual privacy and the belief that any “backdoor” or special access to data for law enforcement would make the world less secure, large tech companies have also staunchly fought back against government efforts to access their encrypted technology. For example, in 2015, the United Kingdom (U.K.) introduced a draft Investigatory Powers Bill. The bill required big tech companies to decrypt communications on the companies' devices on demand as well as retain their users' browser histories for a year so that the U.K. could have access to such information if needed for law enforcement purposes. Apple released statements opposing the legislation. Apple CEO Tim Cook made several public statements opposing the bill, stating, "If you close down the major companies from using encryption, the bad guys aren't going to stop using encryption. They are just going to go to another source." Cook also asserted, "If you leave a back door in the software, there is no such thing as a back door for good guys only .... If there is a back door, anyone can come in the back door.”

In the wake of the Apple v. FBI controversy that unfolded in the spring of 2016 in the U.S., Apple also released statements with equally strident language. Apple's customer letter, within the subheading “Need for Encryption,” contained the following passage:

All that information needs to be protected from hackers and criminals who want to access it, steal it, and use it without our knowledge or permission. Customers expect Apple and other technology companies to do everything in our power to protect their personal information, and at Apple we are deeply committed to safeguarding their data. Compromising the security of our personal information can ultimately put our personal safety at risk. That is why encryption has become so important to all of us. For many years, we have used encryption to protect our customers' personal data because we believe it's the only way to keep
their information safe. We have even put that data out of our own reach, because we believe the contents of your iPhone are none of our business.\textsuperscript{74}

Members of the tech industry commentariat have also decried government efforts to create secure encryption keys.\textsuperscript{75}

Additionally, the Massachusetts Institute of Technology released a 2015 report on end-to-end encryption and the risks of government access titled \textit{Keys Under Doormats}.\textsuperscript{76} The report was written by fifteen computer science and security experts\textsuperscript{77} and is perhaps one of the most comprehensive analyses performed on this subject to date. The experts argued that the price of granting law enforcement exceptional access to user privacy would be more damaging today than it would have been twenty years ago, referring to the Clipper Chip controversy of the 1990s.\textsuperscript{78}

However, not all tech commentators or the general media are in support of end-to-end encryption. A July 18, 2015 editorial by the Editorial Board of the Washington Post requested that Congress heed the points of both sides of the debate when fashioning policy.\textsuperscript{79} The Editorial Board discussed a set of possible limits for Internet freedom centered on the “legitimate needs of U.S. law enforcement.”\textsuperscript{80} The Board also stated that it believes that the Internet should “be subject to the same rule of law and protections that we accept for the rest of society.”\textsuperscript{81}

\textbf{V. THE GLOBAL BATTLE OVER THE REGULATION OF ENCRYPTION}

There does not seem to be a strong correlation between the type of government and the degree of regulation on encryption technology. For example, one of the nations that is proposing one of the strongest sets of regulations is the UK, a country not known in recent history for being authoritarian in its surveillance of its own citizens. Notably, the countries that have typically been known for their libertarian ideals, such as the U.S. and the Netherlands, have wrestled more with the correct course of action.

It appears that the most common thread in the debate over how to regulate encryption is tied to a country's recent experiences with terrorist attacks. Countries where terror has become a larger concern, such as the U.S.,\textsuperscript{82} China,\textsuperscript{83} and France,\textsuperscript{84} have either passed stringent regulations or have debated stringent regulations regarding encryption in recent years.\textsuperscript{85} Conversely, countries surveyed that have not experienced terrorism in recent years have utilized more traditional methods of surveillance to combat Internet crime.\textsuperscript{86}

The following sub-sections discuss differing legislative positions and commentaries from the U.K., China, South Korea and India. The countries are presented by level of government access, from most government access to the least. Following these is a brief discussion of the use of encryption by ISIS, which is the greatest international terrorist concern for countries currently addressing regulation of encrypted information.\textsuperscript{87}

\textit{A. The United Kingdom}

As discussed in Section IV above, the U.K. introduced the draft Investigatory \textsuperscript{364} Powers Bill in 2015.\textsuperscript{88} British officials have explained the legislation's goals, arguing that encrypted messaging systems are acceptable under the new
law provided that the companies can decrypt the messages when asked by the government. The draft bill is not, therefore, an official ban on encryption but is functionally equivalent to a ban on end-to-end encryption.

This draft bill resulted in a wave of lobbying from the tech industry. Apple and other tech giants (including Google and Microsoft) submitted a memorandum to a committee of Parliament which scrutinized the proposed law. Apple has all but admitted that it will be placed in a position where it must comply with overlapping laws, but that it will “be left having to arbitrate between them, knowing that in doing so they might risk sanctions.” Companies have also warned the British government and the public that they will comply if they must, but costs to consumers will inevitably increase. Matthew Hare, a chief executive of the Internet service provider Gigaclear, stated that “the indiscriminate collection of mass data is going to have a massive cost.”

In regards to the U.K. draft Investigatory Powers Bill, Wikipedia founder Jimmy Wales has stated, “[i]t is not possible in any sense of the word for the U.K. to ban encryption. More to the point, it's a moronic thing to do.” Accordingly, “[t]he problem noted by many last year is that a backdoor to encryption, even if euphemistically rebranded as a ‘front door’ or a ‘golden key,’ is by definition a vulnerability. Building in backdoors threatens consumers and makes them vulnerable to criminals and hostile foreign governments alike.” However, Wales has also stated that court-approved warrants would be a reasonable method to deal with Internet crime rather than backdoor access. Golden keys, or secure encryption keys, have been called “mythical nonsense.”

Following the initial public and private response to the bill, it was revised to address some of the tech industry's concerns. Notably, the revised bill confirmed a power the government already retained, which was not part of the original draft. The British government already had the power to “covertly glean personal data for the purposes of ‘preventing death or injury or damage to a person's physical or mental health.’” This was seemingly extended in the revised draft bill by including language allowing access for investigation or prevention of a “serious crime.” Criticism from police chiefs that the bill provided too little power to police had an effect on the revisions as well, specifically in the type of web history that could be requested. The bill originally only allowed police to identify the sender of a message “and where it was suspected an individual accessed illegal material ....” After revision, the draft bill gave the power to ask for any web material that is “necessary and proportionate for a specific investigation.”

The draft bill, if passed, will likely create more disagreement, and possibly litigation, between the British government and the tech industry. A spokesperson of the Home Office, the governmental department in the U.K. responsible for immigration, counter-terrorism, police, drugs policy, and related science and research, was quoted saying that the government wants to “find a way to work with industry” that will ensure “that [the data of] terrorists and criminals [is accessible] in order to resolve police investigations and criminal acts.” Thus, the outcome of the bill will signify the U.K.’s position on encryption and the degree to which it values access over individual privacy concerns expressed by the tech industry.

**B. People's Republic of China**

The National People's Congress (NPC) of China recently developed an antiterrorism law that would require companies to give encryption keys to public officials for any data that is stored on Chinese servers. This would include data flows from domestic and international companies doing business in China.
The law is a reaction to two recent developments. The first is the series of recent attacks carried out by terrorists centered in Xinjiang Uighur Autonomous Province; and the second is the increased attention the Communist Party has been paying to the allegations of Edward Snowden regarding the U.S. government's surveillance programs. The law has many provisions, the first two of which are particularly relevant for discussion: (1) “[b]ackdoors must be made available to government authorities [,]” and (2) “[e]ncryption [k]eys must be made available to government authorities.”

ChinaFile, the online magazine from the Center on U.S.-China Relations, stated that the law would have three major effects on global technology: 1) cyber-sovereignty might further alienate the Chinese Internet from the rest of the Internet; 2) Internet companies will see an increased opportunity to sell products within the Chinese market, but it will likely come at a higher cost; and 3) the international market for Chinese products will decrease. Before the law passed, Scott Livingston, an American attorney specializing in Chinese trade and investment law, stated that “[c]ompanies with operations in China would be wise to familiarize themselves with the present draft, as its content and context may be instructive for understanding the future direction of China's Internet policy.”

Fu Ying, Spokesperson for the Chinese Parliament, stated that the laws will not affect tech companies' reasonable interests. He also stated that many Western governments, including the U.S., have made similar demands. Further, Fu stated that China's use of the encryption data will be more disciplined than American surveillance. He hopes that “with transparent procedures, China's anti-terrorism campaign will be different from what the United States has done: letting the surveillance authorities run amok and turn counterterrorism into paranoid espionage and peeping on its civilians and allies.” This law was passed in December 2015, with its final version excluding some of the most controversial provisions.

C. France

There has been movement towards regulating encryption in France, increasingly so following the Paris attacks in 2015. The National Assembly, France's lower house of Parliament, has led the movement to regulate encryption. A recent amendment to a penal reform bill added extensive punishments for smartphone manufacturers that do not comply with law enforcement requests. The proposed law would require smartphone manufacturers to “provide access to data in connection with terrorism investigations.” If a company does not comply, there is a potential fine of up to $385,000. In addition, executives of a non-complying company could face a potential five-year prison sentence. This amendment (and the bill that it is attached to) would still need to pass the Sénat and be signed into law by the French President (currently President Hollande).

D. India

India does not have an official law or response to the rise of encryption market-wide, but instead has waged individual battles with companies that employ the technology. For example, a blanket law on encryption was passed in 2015 that requires individual citizens to keep the plaintext of any encrypted data they received for up to ninety days, and then to turn the plaintext over to security forces. In response to public backlash, social media apps were exempted from
the law. In 2010, the Indian government nearly banned Blackberry for refusing to give the government access to personal messaging systems. There is much speculation over how the Indian government will handle the recent switch to end-to-end encryption by WhatsApp, a popular messaging system in India.

E. Canada

Canada waged what at least one commentator has called a “quiet war on encryption.” While Canada has not yet spoken on the current end-to-end encryption debate, it is worthwhile to note the country's history with encryption. One significant way in which Canada has waged a “quiet war” on encryption is by widening the type of devices covered by its cryptography standards and weakening an important cryptographic standard. Canadian mobile telecommunications providers must agree to and implement the Solicitor General’s Enforcement Standards (SGES). The standard relevant to this comment is standard twelve. The annotation for the standard reads:

Law enforcement requires that any type of encryption algorithm that is initiated by the service provider must be provided to the law enforcement agency unencrypted. This would include proprietary compression algorithms that are employed in the network. This does not include end to end encryption that can be employed without the service provider's knowledge.

In 2012, the Canadian government expanded the scope of the SGES to include “circuit” forms of communication that were transmitted on new technologies. SGES applies to mobile providers, but successive versions of “lawful access legislation” would theoretically apply to all telecommunications providers. Bill C-13 was passed on December 9, 2014 and contains explicit clauses authorizing authorities to require telecommunications providers to decrypt certain communications. Under Bill C-13, Canadian providers have the authority to challenge the demand.

Canada's foreign signals intelligence agency, the Communications Security Establishment (CSE), has also played an active part in undermining global encryption mechanisms. For example, the CSE ran a multi-national committee at the International Organization for Standardization. This committee led to the NSA propagating and granting a sense of legitimacy to a method of encryption that is known to be vulnerable, the Dual EC DRBG.

F. South Korea

The South Korean government has generally chosen to pursue individual surveillance and subsequent warrants to obtain user data. A South Korean tech company, Daum Kakao, released a transparency report in January 2015 stating that the amount of warrants issued by the courts had quadrupled between 2012 and 2014. Naver, a South Korean search portal, has seen the warrants issued for user data increase six times in the same period. President Park Geun-Hye also recently announced increased surveillance for “any messages deemed as insulting to her or generally rumor-mongering.” This has pushed companies like Daum Kakao to change messaging systems by adding encrypted features. It has also pushed users to use international messaging systems with encrypted features, like Telegram from Germany.
G. The Netherlands

The Netherlands has taken a different approach than most countries in regard to legislation. In the “Government Position on Encryption,” authored by the Minister of Security & Justice and the Minister of Economic Affairs, the Dutch declared that a strongly encrypted Internet was in the country's best interest. Therefore, the Dutch government will not be mandating encryption backdoors of any kind. The Ministers said “[i]t is currently not desirable to take restricting legal measures concerning the development, availability, and use of encryption within the Netherlands.” Also noteworthy was the statement by the Dutch government that backdoors would create “undesirable consequences for the security of communicated and stored information” because “digital systems can become vulnerable to criminals, terrorists and foreign intelligence services.”

H. ISIS

ISIS is known to use the encrypted app Telegram from Germany to communicate amongst its adherents, spread propaganda, and recruit new followers. The terror group continues to use the app even after Telegram took action to shut down several of its channels following the Paris attacks in 2015. Telegram co-founder, Pavel Durov, is a noted privacy rights advocate. Durov relented to government pressure following the Paris attacks and shut down seventy-eight messaging channels used by ISIS.

VI. THE AMERICAN BATTLE OVER ENCRYPTION

The battle over encryption has grown significantly both internationally and domestically in the U.S. in recent years. The U.S. reached a watershed moment in the controversy between Apple and the FBI, when Apple's refusal to comply with the court order of Federal Magistrate Judge Sheri Pym to help decrypt the phone of one of the San Bernardino shooters set off a firestorm of criticism, but also garnered public support from other tech giants such as Google, Microsoft, Facebook, Twitter and Huawei.

Per a Government Status Report filed March 28, 2016 in federal court, the FBI gained entry into the San Bernardino shooter's iPhone without the assistance of Apple, and the case is now closed. However, future U.S. regulatory steps will greatly impact the debate over the role of encryption in American society. This comment argues that the U.S. should defer to the expertise of the tech industry on the dangers of regulating encryption. Allowing back-door access for law enforcement or national security agencies makes systems vulnerable to attacks by hackers and hostile nations. Making U.S. systems more secure while other countries make theirs vulnerable allows the U.S. to be comparatively safer.

A. The United States' Regulatory Options

It seems unlikely that any new legislation overhauling cyber-privacy will do so to the same degree as recent British or Chinese legislation. This can be attributed to several factors. First, domestic political pressures regarding surveillance and privacy are at a boiling point following the Edward Snowden revelations of 2013. Second, requiring access to all encrypted communications through any technological medium would likely stymie innovation in the global tech industry. Third, new overhauling legislation would also increase tension between different markets purchasing
products from these companies, as almost any scheme chosen by the U.S. would likely have a vast precedential effect across the globe.

The dispute between Apple and the FBI points to the likelihood that the U.S. is moving towards creating regulations that provide for more access, but limited to special emergency situations such as terrorist attacks. If the U.S. decides not to regulate, that could lend official legitimacy to the position of the tech industry that *372 strong security systems are essential. If the U.S. were to embrace the previously discussed PrivaTegrity scheme created by David Chaum, that could have far-reaching implications, as well. In examining how the U.S. should best address the regulation of encryption in the future, it is necessary to look at how the U.S. has addressed encryption prior to the current dispute between Apple and the FBI.

**B. Historical Context**

American statutes regarding encryption have often centered around the ability of international criminal and terrorist groups to encrypt data and use that data to hurt the U.S.’ national security and economic interests. Since the 1990s, there have been various regulations on the export of encryption software. Until 1996, this was done by classifying encryption software as “munitions.” Classifying encryption software as such brought it within the purview of the State Department, and was looked at as a “defense article.” The export of encryption, however, is governed by the Arms Export Control Act (AECA) and the Export Administration Act (EAA). The AECA is administered by the State Department under the International Traffic in Arms Regulations (ITAR), and the EAA is administered by the Commerce Department under the Export Administration Regulations. A licensing system, with limited exceptions, primarily enforces these regulations.

American regulation of encryption exports has been challenged multiple times in U.S. courts. In *Bernstein v. United States Department of State*, the petitioner, Daniel Bernstein, sought to publish the source code of encryption software that he had written on the Internet, but he was denied a license by the State Department. Bernstein challenged the application of the AECA and ITAR to the publication of encryption source code, arguing that the denial of a license was a form of prior restraint. Ruling in favor of Bernstein, the court held that source code is speech, which had the effect of loosening export regulations for encryption software. In *Karn v. United States Department of State*, a book of encryption algorithms was ruled also to be speech, but a diskette containing identical information was ruled not to be speech. The diskette instead was ruled to be cryptographic software, and the court ruled that judicial review of the President’s designation of an item as a defense article is barred by AECA § 2778. *Karn* is significant because the District of Columbia (DC) District Court held that AECA and ITAR are constitutional because they furthered an important or substantial government interest. The *Karn* court also held that ITAR does not constitute a prior restraint on free speech, because the regulations are content-neutral.

In 1993, the White House introduced the “Clipper Chip Initiative” as an effort to regulate encryption. The Clipper Chip would have subjected electronic encryption technology to key escrow in order for the government to hold a key to decode messages. Key escrow keeps a record of the key for each device with a Clipper Chip installed. Before two devices with Clipper Chip would communicate, the chips would create a string of data called the “Law Enforcement Access Field” (“LEAF”). The LEAF would provide the government with the digital signature necessary to obtain the keys and decrypt the communication.

---

The Clipper Chip initiative, however, failed due to what has been called “a firestorm of protest on constitutional and economic grounds.” The successful hacking of the Clipper Chip by Matt Blaze also contributed to the Clipper Chip not being adopted by the tech industry. Battles between law enforcement and the tech industry over the Clipper Chip have colloquially been referred to as “The Crypto Wars.”

*374 C. The Current Battle over Encryption

In January of 2015, President Barack Obama stated “if we find evidence of a terrorist plot ... and despite having a phone number, despite having a social media address or email address, we can't penetrate that, that's a problem.” Calling tech companies “patriots,” President Obama said in the same statement that he believes they would like to help solve the problem. These statements were made after meetings with then U.K. Prime Minister David Cameron, with Cameron at Obama's side.

In September of 2015, however, a leaked memorandum from the National Security Council (NSC) contradicted the President's earlier stance. This memorandum was mostly devoid of technical substance and instead posed political solutions. The Administration saw itself with three political options in handling the fight over end-to-end encryption. The first would be to plainly oppose new laws and new backdoors and to speak favorably about encryption. The second would be to defer the issue and consult with experts. The third option would be to “punt the issue into the long grass.” In a surprising turn, however, the memorandum pushes the first option the most. “Overall, the benefits to privacy, civil liberties, and cybersecurity gained from encryption outweigh the broader risks that would have been created by weakening encryption.” The memorandum's position seems to indicate that there are possibly different factions of the Executive Branch of the U.S. government that have diverging stances on encryption.

In October 2015, the Obama Administration stated that they will not attempt to pass new laws regulating end-to-end encryption. In a written statement to the Senate Homeland Security and Governmental Affairs Committee, FBI Director James Comey stated “the administration has decided not to seek a legislative remedy now but that it makes sense to continue the conversations we are having that are very productive.” There is speculation that perhaps the Obama Administration will change course in 2016 following the terrorist attacks in Paris and San Bernardino. Members of the Obama Administration, as well as congressional allies of the Administration, have been increasingly critical of encrypted messaging following the Paris attacks.

1. Legislative Branch Positions

The House Judiciary and the House Energy and Commerce Committee have set up a bipartisan working group to “examine the complicated legal and policy issues surrounding encryption.” The group has been tasked with identifying “potential solutions that preserve the benefits of strong encryption - including the protection of American[s] benefits of strong encryption - while also ensuring law enforcement has the tools needed to keep us safe and prevent crime.” The group, according to its own press release, has jurisdiction over the entire nexus between these issues.

Senator Richard Burr (R-N.C.) and Senator Dianne Feinstein (D-Calif.), the chairman and vice chairwoman of the Senate Intelligence Committee, introduced legislation on April 13, 2016 that “mandate[s] that companies assist law
enforcement in accessing content shielded by encryption.”

There is also draft legislation to form a commission to examine the effect of encryption on national security. Rep. Michael McCaul (R-Tex.), the chairman of the House of Representatives Homeland Security Committee, and Senate Intelligence Committee member Mark Warner (D-Va.) have drafted legislation to establish the commission. McCaul plans for the commission to be comprised of “technologist[s], privacy activists, academics, and law enforcement official[s].” Also planning to have input in the commission is Representative Will Hurd (R-Tex.), who worked prior to his congressional career for a cybersecurity firm. This legislation is considered more likely to become law than the Feinstein-Burr proposal, possibly because it is more moderate in its approach.

Senator Chuck Grassley (R-Iowa), the chairman of the Senate Judiciary Committee, has also sent a letter to Deputy Attorney General Sally Yates in which he voiced support for Comey’s positions. “Countries like Great Britain and France are much further along in their national dialogues about how best to balance privacy and public safety with regard to encryption and are currently contemplating specific legislative proposals to address the threat posed by widespread inviolable encryption.”

Following the Paris attacks, Senator John McCain (R-Ariz.) publicly expressed support for legislating encryption. Senator McCain stated: “In the Senate Armed Services we’re going to have hearings on it and we’re going to have legislation.” He also stated that the non-legislative status quo is “unacceptable.” Former Representative Mike Rogers (R-Mich.), once the chairman of the House Permanent Select Committee on Intelligence, also recently publicly criticized moves towards end-to-end encryption. In an op-ed for CNN’s website, Rogers states that encryption technology poses a significant security risk to the U.S. Rogers argues it does this by allowing radicalized individuals to communicate in a manner that is not accessible either by law enforcement or by the companies that are facilitating the communication. Rogers posits that the American tech industry has made sacrifices to its bottom line before in the name of the public good, and the tech industry acquiescing on end-to-end encryption would keep the public safe.

2. U.S. Law Enforcement Positions

The majority of domestic anti-encryption rhetoric comes from U.S. law enforcement institutions. The most vocal opponent of end-to-end encryption is James Comey, Director of the FBI. Director Comey argues that end-to-end encryption forces the FBI to “go dark” on investigations. Director Comey has stated that there are “societal costs to universal encryption,” and that if law enforcement cannot access information and save lives, the nation would be in a “very dark place.” Director Comey’s solution would create a specific passcode or key to decode encrypted messages that would only be available to law enforcement.

William Bratton, the former Commissioner of the New York City Police Department (NYPD) also critiqued end-to-end encryption and encrypted communications in general, arguing that encrypted communications have made it “impossible for officials to collect warnings on terrorist attacks.” Commissioner Bratton takes his critique one step further, effectively alleging that tech companies profit from the commercialization of secrecy: “We, in many respects, have gone blind as a result of the commercialization and the selling of these devices, that cannot be accessed either by the manufacturer or, more importantly, by us in law enforcement, even equipped with search warrants and judicial authority.”
The National Security Agency (NSA) and its parent agency, the Department of Defense (DOD), have released conflicting statements regarding encryption. Statements from the CIA further muddle a coherent stance on encryption. The NSA's statements are similar to statements made by the FBI, voicing concerns about “going dark” and how that will affect the agency's ability to properly fulfill its mission. NSA Director Admiral Michael S. Rogers has suggested requiring companies to create a digital key for certain devices but dividing the key into pieces so that interpersonal or interagency cooperation would be needed to use it.

Officials from the DOD, however, have recently supported end-to-end encryption. In an interview, security expert and former Vice Chairman of the Joint Chiefs of Staff Admiral James A. Winnefeld responded to a question by stating: “[B]ut I think we all win if our networks are more secure. And I think I would rather live on the side of secure networks and a harder problem for Mike on the intelligence side [speaking of Adm. Michael S. Rogers of the NSA] than very vulnerable networks and an easy problem for Mike.”

In a speech on encryption on November 16, 2015 at the Center for Strategic and International Studies, CIA Director John Brennan provided statements which seemed to put the CIA's relative views on privacy and national security into perspective, hoping that the recent Paris attacks would be a “wake-up call.” Former CIA Deputy Director Michael Morell has also stated that, at the very least, the Paris attacks will renew the American discussion over end-to-end encryption.

VII. THE U.S. SHOULD DEFER FROM CRAFTING NEW ENCRYPTION REGULATIONS

The global debate over end-to-end encryption is complex and hard to navigate. One of the primary reasons that it is difficult to form any sort of nuanced understanding or coherent policy proposal is because there are so many different requirements amongst varied nations. The varied standards are a problem because, according to tech experts, any amount of access can make a system more vulnerable. However, this comment argues that the U.S. should hold off on crafting new regulations, as being patient throughout the duration of this debate will make the U.S. comparatively safer than other nations. Full encryption makes us safer, and while other countries are making their own systems more vulnerable, we should be strengthening our communication systems instead.

Smartphone data has undoubtedly been helpful in obtaining information and capturing criminals. According to François Molins, the Chief Prosecutor of Paris, smartphone data was critical in investigating those responsible for the attacks on Charlie Hebdo magazine on January 7, 2015. Molins also has stated that it was integral into the investigation of an attack on a gas station in Saint-Quentin-Fallavier in June of 2015. As authors of a recent op-ed in The New York Times argued, in at least seventy-four legal proceedings in Manhattan between October of 2013 and June of 2014, iPhones were inaccessible because of the phones' encryption. The iPhones were connected to investigations of child sexual abuse, sex trafficking, and many other violent crimes. Following the recent court order served upon Apple by Judge Pym, Cyrus Vance, Jr., District Attorney of Manhattan, stated that he has 175 phones currently waiting to be unlocked. Access to these phones could allow law enforcement to prosecute the perpetrators of these crimes and prevent further criminal acts.
The benefits of end-to-end encryption, however, outweigh the possible security drawbacks. A completely secure network is more secure than a partially secure network, even if the part that is unsecure is small. It is better to fight crime with the tools we have than to open the floodgates. The Intercept, an online journal seeking to hold governments and corporations accountable, 238 has claimed that many of the statistics and anecdotes used by law enforcement to justify advocating for stronger encryption controls are not factually correct. 239 Citing a Federal Courts report on wiretapping, The Intercept stated that in only four cases were state and federal police obstructed by encryption. 240 In a direct investigation into the claims made by FBI Director Comey, The Intercept found that Comey had actually misconstrued several stories that he told in speeches supporting encryption controls. 241

Providing access to encrypted data for law enforcement agencies and other governmental bodies across the world makes all related encrypted systems more vulnerable, which in turn makes the global population more susceptible to cyber-attacks and hacking. 242 The previously mentioned Keys Under Doormats report thoroughly explains the technical risks associated with governments weakening encryption in the name of law enforcement. 243 The report lists three reasons why regulating encryption results in systems is less safe than what exists currently. 244 The first is that providing exceptional access to these communications would force tech companies to make a “U-turn from the best practices now being deployed to make the Internet more secure.” 245 Tech companies have been increasingly utilizing “forward secrecy,” which requires that transaction keys are discarded after every transaction in which they are used. 246 Forward secrecy means that hackers have less to work with, and historical data is safe. 247 Requiring exceptional access would make data more vulnerable because if the access key was hacked, there would be much more information for the potential hacker to recover. 248 Allowing more information to be vulnerable to a hack is certainly an important risk to consider.

The second concern expressed in Keys Under Doormats is that access would increase complexity, and therefore increase vulnerability. 249 The report begins this analysis with the premise that if “exceptional access” (government access to encrypted information) were to be applied throughout the industry, it would create “widespread exceptional access,” meaning use beyond the government. 250 While 249 250 251 this is not the position this comment takes, the consequences the report lists are still important to consider. To achieve access, new technologies would need to be deployed and tested, which would increase a system's vulnerabilities. As Dr. Frederick R. Chang, the former head of research for the NSA, testified before the U.S. Congress in 2013, “[i]ndeed it has been said that complexity is the enemy of security.” 252 To allow our encryption systems to become exponentially more complex is to allow for an increase in vulnerabilities. This is an unacceptable risk.

The final risk discussed by Keys Under Doormats is that allowing exceptional access for law enforcement will put a focus on who bad actors should aim to attack. 253 The report cites the hacking of the Office of Personnel Management (OPM) of 2015 as an example of “how much harm can arise when many organizations rely on a single institution that itself has security vulnerabilities.” 254 This is the risk given by the report that is the most persuasive. To allow for exceptional access is to trust our most complex technological systems to the hands and minds of those that are not the most technologically capable.

Bruce Schneier, a board member of the Electronic Frontier Foundation (EFF) and noted tech commentator, 255 has stated that many of the most noted tech security failures of recent history are correlated to failures in encryption technology. 256 These failures are evidenced in such instances as the Chinese government’s hack of the OPM and the hacks of the TJX Companies and Target Corporation. 257 Clearly, a hack akin to the OPM hack affects our national...
security, as the names and personal information of government workers are involved. However, private corporations that are hacked can also affect our national security. These hacks could be training for later attacks on federal or other government entities. Also, federal employees and agencies are likely regular customers with these types of companies, and any type of information could be found and exploited in a hack. With stronger encryption, or at least less vulnerable networks, these hacks would likely be much more difficult. As Nuala O'Connor stated, “the sophistication and number of cyber attacks on the security of individuals, businesses, government agencies, and critical infrastructure are only increasing.”

Apple v. FBI also highlighted a third concern stated in Keys Under Doormats—that to concentrate focus on a specific set of actors is to introduce an unacceptable risk due to the possibility that one of those actors could be compromised. Focusing on a specific set of actors would draw attention and create an enlarged risk due to exceptional access; a third party, such as law enforcement, would have security access to all actors of interest. If one of the parties of interest was attacked, there is an increased potential that the attacker would be able to access other associated actors' information since they could access the keys to all associated actors through a third party.

There is also an argument that “going dark” is the wrong metaphor, and that encryption does not significantly hinder law enforcement's legitimate investigative tactics. The authors of a report called Don't Panic: Making Progress on the “Going Dark” Debate, published by the Berkman Center for Internet & Society at Harvard University, recently argued this point. The authors of Don't Panic state that due to several factors, the gaps in information left by encryption will be filled. This is outside the scope of this comment, but it is worth noting that according to this paper's thesis, encryption does not make us less safe, and furthermore, there are opportunities to fill in the gaps by other measures. This argument is supported by the Chertoff Group's Ground Truth white paper, in its assertion that “we can find no successful terrorist attack that would have been prevented by the availability of a lawful decryption access technology.”

As explained in Keys Under Doormats, and expounded upon by Mr. Schneier and the Motherboard article, changing course and demanding access to encryption would make personal information less safe. Many daily activities utilize encryption technology, such as online banking, shopping, and emails (personal and professional). These sensitive sets of information are best served by the most secure systems possible. And when viewed with the perspective of the intrusive regulations and laws being passed in other nations, allowing for strong encryption would make us comparatively safer because our personal information would be secure.

As a result of the already changed or currently changing legislation in other nations, the U.S. finds itself in a unique and potentially advantageous position. By not acting to regulate end-to-end encryption, the U.S. could actually make itself comparatively safer than other nations. Other countries have made their citizens more vulnerable to hacking or cyber-attacks as a result of the countries' regulations. By not regulating, the U.S. retains its level of security and stature in the global tech market while simultaneously becoming more secure. Ground Truth quotes General Michael Hayden as saying, “[b]ecause of the nature of global telecommunications, we are playing with a tremendous home-field advantage, and we need to exploit that edge ....”

Having this built in “home-field” advantage over tech companies should be a boon to the American economy and national security. It would keep capital and innovation within American borders. In addition, the metadata that is associated with the encrypted data that passes through the U.S. will likely continue to do so if encryption remains strong. For example, Brazil has announced that the nation plans to build Internet cables that go under the Atlantic Ocean to
reach Portugal. Brazil plans to do this to avoid U.S. regulations on data collection. If the U.S. regulates encryption, it would take away both the capital associated with the existing Brazilian communications passing through the U.S. as well as the metadata associated with the encrypted data. Thus, if the U.S. wants to maintain its “home-field” advantage, it should not pass heavy-handed regulation of encryption technology.

VIII. CONCLUSION

The state of confusion created around the issue of encryption is a result of diverging methods of regulation around the world and differing views of the utility of the technology within the U.S. government. All of these divergent stances and policies ultimately make global networks less secure and thus make American information and data less safe. However, the U.S. is in a unique position as a global leader in both innovation and in the transportation of communication. Encryption makes data more secure and makes countries safer from hackers and terrorists. The U.S. should defer to the tech industry's knowledge on encryption, as doing so will make its data and its communication systems comparatively more secure, which will increase personal and national security. The creators of these encrypted systems know the consequences of compromising the systems' integrity better than anyone else and the relevant actors concerned with this issue should defer to their positions. Choosing between security of encrypted networks and an efficient and complete effort against crime and terror is a false choice.

Footnotes

1. J.D., Temple University Beasley School of Law, 2016. I would like to thank Associate Dean Duncan Hollis for his feedback and guidance throughout writing this, as well as the Editorial Board and Staff Editors that have improved this article throughout the last year. I also want to thank my family and friends for their patience and support throughout law school. Lastly, thank you Liz, for your encouragement to write this, and for everything.

2. See Anusha Asif, Apple vs. FBI: Encryption Case Timeline, TECH NEWS TODAY.COM (Feb. 29, 2016, 7:07 AM), http://www.technewstoday.com/28773-apple-vs-fbi-encryption-case-timeline/ (discussing the controversy between Apple and the FBI, where the FBI attempted to enforce a court order forcing Apple to de-encrypt the phone owned by one of the shooters who committed a mass shooting and attempted bombing in San Bernardino, California).


6. Telegram is a messaging and social networking application allowing users a wealth of options, including state-of -the-art encryption, a relatively private tool for dissemination to large groups, and which is, most importantly, free. Jamie Dettmer, Why Islamic State Loves Telegram, VOICE OF AMERICA (Jan. 8, 2008, 1:55 PM), http://www.voanews.com/content/why-islamic-state-loves-telegram/3137040.html. There are many different names for the group ISIS. Faisal Irshaid, Isis, Isil, IS or D'aesh? One Group, Many Names, BBC NEWS (Dec. 2, 2015), http://www.bbc.com/news/world-middle-east-27994277. For the purpose of this comment, the group is referred to as ISIS, unless called ISIL or Daesh by the source referenced.

7. See Harold Abelson et al., Keys Under Doormats: Mandating Insecurity by Requiring Government Access to All Data and Communications, COMPUTER SCIENCE & ARTIFICIAL INTELLIGENCE 11 (July 6, 2015) (examining the use of encryption in popular social media platforms today); see infra Section II (further defining public and private keys).
See infra Section II (examining the use of encryption in popular social media platforms today).

See Abelson et al., supra note 6, at 11 (explaining that transformed data was protected by a symmetric key that operates in a mode where the sent data was encrypted and can only be accessed by the receiver); see also Plain text, WEBOPEDIA, http://www.webopedia.com/TERM/P/plain_text.html (last visited Oct. 23, 2016) (“In cryptography, plain text refers to any message that is not encrypted.”). Plain text, plain-text, or plaintext is also used to refer to any text that contains only text and does not support formatting, including basic formatting such as italicization. Plain text, COMPUTER HOPE, http://www.computerhope.com/jargon/p/plaintext.htm (last visited Oct. 23, 2016).


In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. ED 15-0451M (C.D. Cal. Feb. 16, 2016).

See Kalia, supra note 10 (arguing that the government intended to enforce backdoors, an extra-legal process, to attempt to obtain user data from technology companies).

Id.


Saunders, supra note 15, at 948.

Id.

Id.

Id.

Id. Public key encryption can also generate digital signatures, which authenticate the identity of the sender of a message and also make the content of the message available. Id. at 948-949.


belief that encrypted technology was used in the attack and indicating that because the terrorists used widely available encryption tools to communicate with each other, it was difficult for investigators to monitor their conversations), with Jeff Larson & Julia Angwin, Fact Checking the Debate on Encryption, PROPUBLICA (Dec. 15, 2015), https://www.propublica.org/article/fact-checking-the-debate-on-encryption (placing the possible use of an encrypted telephone by one of the Paris attackers in a larger context).


See id. (stating that 130 people were killed); see also Adam Chandler, Krishnadev Calamur, & Matt Ford, The Paris Attacks: The Latest, THE ATLANTIC (Nov. 22, 2015), http://www.theatlantic.com/international/archive/2015/11/paris-attacks/415953/ (stating that 368 people were injured).


See Ellen Nakashima, Apple Vows to Resist FBI Demand to Crack iPhone Linked to San Bernadino Attacks, WASH. POST (Feb. 17, 2016), https://www.washingtonpost.com/worldnational-security/us-wants-apple-to-help-unlock-iphone-used-by-san-bernardinoshooter/2016/02/16/69b903ee-d4d9-11e5-9823-02b905009f99_story.html (explaining the controversy between Apple and the FBI over an iPhone used by one of the San Bernadino shooters).


Id.

Id.

See Asif, supra note 1 (recognizing that no “backdoor” exists for Apple's product and therefore administrators needed to create one to see encrypted information).


See Andy Greenberg, Hacker Lexicon: What Is the Dark Web, WIRED (Nov. 19, 2014, 7:15 AM), http://www.wired.com/2014/11/hacker-lexicon-whats-dark-web/ (“Even Facebook has launched a Dark Web site aimed at better catering to users who visit the site using Tor to evade surveillance and censorship”). Tor and the Dark Web are very complex issues, so this will be a relatively basic explanation of both. Based out of a nonprofit that does research and development into Internet privacy, Tor offers a technology that bounces users’ and websites’ traffic through thousands of “relays” provided by volunteers across the globe. This makes tracking the source of the information or the location of the user difficult. Tor was originally created by the U.S. Navy and received much of its initial funding from the State Department. Tor: Overview, TOR PROJECT, https://www.torproject.org/about/overview (last visited Oct. 20, 2016). Tor is mainly used for lawful purposes by five different groups: (1) normal people who want to keep their internet activities private from websites...
and advertisers, (2) those concerned about cyberspying, (3) users evading censorship present in their part of the world, (4) military personnel, and (5) activists and journalists evading censorship. Tor is also used by criminals and terrorists across the world to maintain anonymity. Criminal users use Tor and similar anonymity networks on what is called the “Dark Web”. See Who Uses Tor?, TOR PROJECT, https://www.torproject.org/about/torusers.html (last visited Oct. 20, 2016) (describing different groups that use Tor as well as their objectives). The Dark Web is a subset of the Deep Web, which is essentially the area search engines cannot crawl for or index. The Dark Web and the Deep Web are terms that are often used interchangeably.


There is debate on the integrity of the end-to-end encryption used by WhatsApp. An employee of the Dutch Intelligence Service AIVD, speaking during a seminar marking the AIVD's 70th anniversary, stated that “I would not trust that WhatsApp is so safe.” Janene Pieters, Dutch Intelligence Service Warns of WhatsApp Security Issue, NLTIMES.NL (Nov. 6, 2015, 12:27 PM), http://www.nltimes.nl/2015/11/06/dutch-intelligence-service-warns-of-whatsapp-security-issue/. This was in response to a question regarding whether or not AIVD officers could read communications taking place over WhatsApp. Id.

Klint Finley, New Facebook Feature Shows Actual Respect for Your Privacy, WIRED (June 1, 2015, 5:42 PM), http://www.wired.com/2015/06/new-facebook-feature-shows-actual-respect-privacy/.

See id. (discussing Facebook’s attempt to strengthen users’ privacy by incorporating encryption into their service, including the encryption of email notifications sent by Facebook).

See infra Part IV (discussing the position of major tech industries of allowing a special key to break encryption).

See infra Part VI (discussing the aftereffects of the watershed controversy between Apple and the FBI).

See infra Section VI.C.2 (arguing that only certain government agencies in certain nation states should have access to encrypted communications and largely only in exceptional situations).

Andy Greenberg, The Father of Online Anonymity Has a Plan to End the Crypto War, WIRED (Jan. 6, 2016, 7:00 AM), http://www.wired.com/2016/01/david-chaum-father-of-online-anonymity-plan-to-end-the-crypto-wars/. David Chaum is the inventor of many encryption protocols and founder of the International Association for Cryptologic Research (IACR). His idea would work practically as a smartphone application. The application, still in its beta testing, is meant by Chaum to be more secure and anonymous than the Tor Network. However, it would have an intentional backdoor that would allow for the privacy and anonymity to be removed from someone “generally recognized as evil.” Id. The insurance for the security of the system comes from the backdoor being divided amongst many key-holders in many different and diverging jurisdictions. Id.

See infra Part IV (discussing the position of major tech industries of allowing a special key to break encryption).

Id.
See Section V.G (explaining that the Netherlands would not mandate any encryption backdoors because of the belief that a strongly encrypted Internet was in the country’s best interest).

Sarah Jeong, A ‘Golden Key’ for Encryption is Mythical Nonsense, MOTHERBOARD (July 21, 2015, 8:45 AM), http://motherboard.vice.com/read/a-golden-key-for-encryption-is-mythical-nonsense.

See Cook, supra note 3 (stating the necessity of encryption as well as Apple's opposition to government regulation because of the threat on data security and personal safety).


See, e.g., Benjamin Wittes, Thoughts on Encryption and Going Dark, Part II: The Debate on the Merits, LAWFARE BLOG (July 12, 2015), https://www.lawfareblog.com/thoughts-encryption-and-going-dark-part-ii-debate merits; see also Editorial Board, Putting the digital keys to unlock data out of reach of authorities, WASH. POST (July 18, 2015), https://www.washingtonpost.com/opinions/putting-the-digital-keys-to-unlock-data-out-of-reach-of-authorities/2015/07/18/d6aa7970-2beb-11e5-a250-42bd8126cf09_story.html (discussing the editorial about the dangers of full encryption and the balance needed between law enforcement and the tech industry) [hereinafter Editorial Board]; see infra Sections V.A.-D (explaining that the governments of the UK, the People's Republic of China, France, and India have each come up with their own policies in regard to regulation of encryption technologies).


See Vance et al., supra note 51 (citing to chief prosecutors from multiple major international jurisdictions).

Id.

See infra Section VI.C.1 (further discussing the view that secure systems are safer, but access is sometimes necessary).


See id. (“Countries including the United Kingdom, Australia, and China have passed or are contemplating laws seeking government access to communications similar to that sought by U.S. authorities.”); see also Nicole Perlroth, Security Experts Oppose Government Access to Encrypted Communication, N.Y. TIMES (July 7, 2015) (discussing the government proposal for “exceptional access” to encrypted data).

Greenberg, supra note 44.

Id.

Id.

See Sophie Curtis, Wikipedia Founder Urges Apple to Stop Selling iPhones in UK If Government Bans Encryption, TELEGRAPH (Nov. 4, 2015, 1:29 AM), http://www.telegraph.co.uk/technology/jimmy-wales/11974687/Wikipedia-
Founder-urges-Apple-to-stop-selling-iPhones-in-UK-if-government-bans-encryption.html (indicating that Apple, Facebook, Snapchat, WhatsApp, Google, and other communication providers use end-to-end encryption to protect user data).

63 See Abelson et al., supra note 6 (concluding that costs associated with creating a special key would be substantial and greatly damage innovation); see also Asif, supra note 1 (discussing Apple's refusal to create a backdoor key for use by the FBI).

64 See Curtis, supra note 6 (providing a quote from Nigel Hawthorn, a spokesperson for a cloud security company, indicating that the government misunderstands how end-to-end encryption works).

65 Cook, supra note 3; see also Editorial Board, supra note 52 (indicating that hardware and software industry associations are opposed to measures that would undermine encryption technology).

66 See Nakashima & Gellman, supra note 57 (discussing how Yahoo's chief of information security believes that law enforcement is requesting security vulnerabilities that hackers and foreign spy agencies could exploit).

67 See Cook, supra note 3 (explaining customer expectation of privacy and protection in technology products); see also Chris Smith, New U.K. Law Might Force Apple to Decrypt the iPhone for Police Investigations, BGR (Nov. 3, 2015, 10:00 PM), http://bgr.com/2015/11/03/iphone-encryption-u-k-law/ (“These companies' reputations rest on their ability to protect their users' data.”).

68 Smith, supra note 67.

69 Id.


71 Id.

72 Id. This position has been seconded by Nigel Hawthorn, spokesperson for cloud-based security company Skyhigh Networks: “‘There's a complete misunderstanding of how end-to-end encryption works. It's wrong to assume that forcing technology companies to break their own security is going to please the average man on the street, and this is not even technically possible in many instances ... despite the inevitable backlash from technology experts, politicians continue to announce these ill-thought-out unworkable proposals.’” Curtis, supra note 62.

73 See, e.g., Cook, supra note 3.

74 Id.

75 Jeong, supra note 48.

76 Abelson et al., supra note 6.

77 Id.

78 Id. For more information about the Clipper Chip controversy see infra Section VI.B.

79 Editorial Board, supra note 52.

80 Id.

81 Id.


84 See Assemblée Nationale, Amendement No. 92CL (France) (introducing an amendment following the spate of 2015-2016 terrorist attacks in France).

85 Id.; Selyukh & Henn, supra note 82; Livingston, supra note 83.

86 See infra Sections V.F., V.G. (explaining that South Korea and the Netherlands have not weakened encryption and will use warrants to obtain data). As this comment does not have an exhaustive survey of countries, this analogy has the potential to be flawed.

87 See infra Sections V.H. (discussing the threat posed by ISIS members utilizing encryption technology).

88 Smith, supra note 67.


90 Id.


92 Id.


94 Id.


96 Id.

97 Id. However, this seems to be inherently contradictory to Wales’ own statements in the same article.

98 Jeong, supra note 48.


100 Id.

101 Id.

102 Id.

103 Id.

104 Burgess, supra note 89.

105 Livingston, supra note 83.
Significantly, the draft of this law does not require that the government give notice when requesting a backdoor.  

Livingston, supra note 83. Articles 15 and 16 deal with this requirement, most relevant to this comment. Article 15 requires that Internet service providers “report their encryption scheme” to “departments responsible for encryption for examination” (likely the Office of State Commercial Cryptography Administration). Id. Article 16 “requires that ISPs that provide encrypted transmission services ... file their encryption scheme with network communication departments and public security organs ....” Id. Article 16 also requires that ISP’s subsequently help in any investigation that results from the inspection of an encryption scheme. Id.

See id. (indicating that Chinese telecom equipment manufacturers Huawei and ZTE Corp. have been effectively kept out of the U.S. market due to cybersecurity concerns).

See Shannon Tiezzi, China’s New Anti-Terrorism Law, THE DIPLOMAT (Dec. 29, 2015), http://thediplomat.com/2015/12/chinas-new-anti-terrorism-law/ (stating that the statute provisions requiring companies to store user data on servers within China and to allow the Chinese government to review their encryption systems were replaced).


See id. (discussing various proposals for reform with the National Assembly). Specifically, the push for penalizing smartphone makers has come from the right-wing opposition faction within the National Assembly. Id.

There has been debate over the efficacy of this amendment in stopping the types of communications that it has been presumably drafted to stop. The amendment would punish smartphone manufacturers, but not the makers of widely-used applications that use encrypted technology (Facebook, Viber, etc.). Reisinger argues that providing a backdoor into hardware might grant some access (SMS messaging, Internet searches, etc.), but that applications could still be used to communicate, leaving law enforcement in the dark. Id. For the text of the amendment, see Assemblée Nationale, Amendement No. CL92 (France).

Reisinger, supra note 118.

Id.

Id.

Id. A deal was eventually struck between Blackberry and the Indian government that allowed the monitoring of Blackberry messaging and emails sent on the devices. *Id.*

*See id.* (discussing Indian encryption laws in the context of WhatsApp's switch to end-to-end encryption).


*See id* (indicating that the Canadian government has cultivated an inadequate security standard such that Canadians are unable to communicate securely over voice or SMS messaging, unless using a third-party application).

Id.

“Standard twelve states, ‘[i]f network operators/service providers initiate encoding, compression or encryption of telecommunications traffic, law enforcement agencies require the network operators/service providers to provide intercepted communications en clair.’” *Id.* There are 22 standards. *Id.*

Id.

Parsons & Israel, *supra* note 130. Examples of circuit-based communications are SMS messaging, MMS messaging, fax, and voice-based communications. *Id.*

Id.

Id.

Id.

The CSE is Canada's primary cryptographic organization and its U.S. counterpart is the National Security Agency (NSA). The CSE has a history of intentionally “providing weak encryption to potential intelligence targets.” *Id.*

Id.

Parsons & Israel, *supra* note 130. The Dual EC DRBG is the Dual Elliptic Curve Deterministic Random Bit Generator.


Kim, *supra* note 142.

145 Brandom, supra note 142.

146 Kim, supra note 142.

147 Brandom, supra note 142.


149 Id.

150 Id.


152 See Pamela Engel, ISIS Has Figured Out Ways to Get Around Restrictions on One of the Main Apps It Uses for Propaganda, BUS. INSIDER (Nov. 24, 2015, 4:46 PM), http://www.businessinsider.com/isis-telegram-channels-2015-11 (describing the finding by Telegram of channel feature used by ISIS to circulate unlimited number of messages to thousands of followers).

153 Id.


156 Asif, supra note 1. Opposing Apple's refusal to comply with the government's request, Donald Trump argued that consumers should boycott the company's products until it agreed to assist the authorities in their investigation. Similarly, the Department of Justice accused Apple of refusing simply as a means of receiving public exposure. Id.

157 Ellen Nakashima, Apple Vows to Resist FBI Demand to Crack iPhone Linked to San Bernardino Attacks, WASH. POST (Feb. 17, 2016), https://www.washingtonpost.com/world/national-security/us-wants-apple-to-help-unlock-iphone-used-by-san-bernardino-shooter/2016/02/16/69b903ee-d4d9-11e5-9823-02b905009f99_story.html; see also Asif, supra note 1 (listing prominent tech companies that stood behind Apple in its refusal to provide a backdoor to the government partly due to the fear of the precedent this case will set.).

158 Selyukh, supra note 4.


160 See NSA Leaks: A Timeline, AL JAZEEBA AMERICA, http://america.aljazeera.com/watch/shows/fault-lines/FaultLinesBlog/2013/11/1/nsa-leaks-a-timeline.html (last visited Oct. 21, 2016) (citing various headlines since Snowden's leak concerning NSA surveillance and data privacy). Snowden was responsible for leaking classified National Security Agency (NSA) documents to the public. For example, the NSA's spying capabilities pertaining to each country were revealed, as well as the fact that the agency had collected nearly 3 billion pieces of intelligence on U.S. citizens. Id.


162 See Saunders, supra note 15 (“The growing use of encryption has led to concerns on the part of federal law enforcement and national security authorities that these technologies will be employed for criminal and terrorist purposes.”).

Saunders, supra note 15, at 949-950.


Id. The United States District Court for the Northern District of California, as well as the United States Court of Appeals for the Ninth Circuit, ruled that the petitioner was engaging in “speech” by publishing their source code, and that the regulation therefore violated the First Amendment. Id.

Saunders, supra note 15, at 951-952.


Id. at 6.

Saunders, supra note 15, at 952.

Id.


Id.

Id.

Saunders, supra note 15, at 952.


Greenberg, supra note 44.


Id.


McCarthy, *supra* note 185.

Id.

Id.

Id.

Id.


Id.


Id.

Id.


Id.


Id.

Id.

See Nasr, *supra* note 200 (discussing the likelihood of either the Feinstein-Burr or the McCaul-Warner bills becoming law).


Id.

Id.

Id.


Id.

Id.

See id. (indicating that tech companies can balance their bottom line with public safety).


Lorenzo Franceschi-Bicchierai, *Encryption Is 'Depressing,' the FBI Says*, MOTHERBOARD (May 21, 2015, 12:37 PM), http://motherboard.vice.com/read/encryption-is-depressing-the-fbi-says. “Going dark” refers to not being able to access key information during an investigation. Id.

Id.

Id.

Id.

See J. David Goodman, *New York City Police Commissioner Says Attacks Will Force Changes in Tactics*, N.Y. TIMES (Nov. 16, 2015, 1:38 PM), http://www.nytimes.com/live/paris-attacks-live-updates/bratton-says-attacks-will-force-law-enforcement-to-change-tactics (“This is something that is going to need to be debated very quickly because we cannot continue operating where we are blind in the area of gathering intelligence on potential attacks”); see also Sanger & Perlroth, *supra* note 50 (noting Bratton's assertion that officials are blinded by encryption).


See Selyukh, *supra* note 196 (indicating how CIA Director Brennan viewed the Paris attacks as a “wake-up call” to the need to reform encryption policy); see also infra note 219 and accompanying text (indicating that the Paris attacks should be seen as a wake-up call for encryption reform).

See Nakashima & Gellman, *supra* note 57 (noting that requiring the agency to use “brute-force” methods is time-consuming and that obtaining covert access to manufacturers requires an often unavailable level of specialty).

Id.
Cushing, supra note 221.

Id.

See Selyukh, supra note 196 (indicating the Paris attacks would be a wake-up call for increasing ways to deal with encryption). Brennan stated: “There has been a significant increase in the operational security of a number of these operatives and terrorist networks as they’ve gone to school on what it is that they need to do in order to keep their activities concealed from the authorities. And as I mentioned, there are a lot of technological capabilities that are available right now that make it exceptionally difficult both technically as well as legally for intelligence security services to have the insight they need to uncover it.” Brennan Delivers Remarks at the Center for Strategic & International Studies Global Security Forum 2015, Central Illigence Agency (Nov. 16, 2015), https://www.cia.gov/news-information/speeches-testimony/2015-speeches-testimony/brennan-remarks-at-csis-global-security-forum-2015.html. Brennan continued: “In the past few years because of a number of unauthorized disclosures and a lot of hand-wrining over the government's role in the effort to try to uncover these terrorists, there have been some policy and legal and other actions that are taken that make our ability, collectively, internationally, to find these terrorists much more challenging. And I do hope that this is going to be a wake-up call.” Id.


See, e.g., Jeong, supra note 48.

See Vance et al., supra note 51 (discussing recent shootings and smartphone security).

Id.

Id.

Id.

Id.


See McLaughlin, supra note 53 (claiming that the examples the FBI relied upon to justify weakening encryption had nothing to do with encryption).

Id.

See Dan Froomkin & Natasha Vargas-Cooper, The FBI Director's Evidence Against Encryption is Pathetic, INTERCEPT (Oct. 17, 2014, 12:49 PM), https://theintercept.com/2014/10/17/draft-two-cases-cited-fbi-dude-dumb-dumb/ (finding that encryption had not been a barrier to identifying or capturing culprits, despite Comey's assertion that encryption could lead law enforcement to miss crucial evidence).

See Jeong, supra note 48 (noting that building any backdoors threatens consumers); see also Cushing, supra note 221 (noting that DOD supports end-to-end encryption).
See Abelson et al., supra note 6, at 15 (stating that the Keys Under Doormats report was a comprehensive report on end-to-end encryption and the risks of government access released in 2015 by MIT).

Id. at 2.

Id.

Id. at 12.

Id.

Id.

Abelson et al., supra note 6, at 2.

Id.

Id.

Id. at 16.

Id. at 2.

See id. (explaining that the attack on the OPM caused multiple federal agencies to lose important data because they used a single organization, OPM, which had an infrastructure security breach); see also Ellen Nakashima, Hacks of OPM Databases Compromised 22.1 Million People, Federal Authorities Say, WASH. POST (July 9, 2015), https://www.washingtonpost.com/news/federal-eye/wp/2015/07/09/hack-of-security-clearance-system-affected-21-5-million-people-federal-authorities-say/ (“Two major breaches last year of U.S. government databases holding personnel records and security-clearance files exposed sensitive information about at least 22.1 million people, including not only federal employees and contractors but their families and friends.”).

Schneier is also a fellow at Harvard's Berkman Institute, the Chief Technology Officer of Resilient Systems, and is regularly published for his commentary on security issues. See, e.g., Bruce Schneier, About Me, SCHNEIER.COM, https://www.schneier.com/ (last visited Oct. 26, 2016).


Id.

See Abelson et al., supra note 6, at 2 (giving an example of how, when the OPM was attacked, it caused harm to other connected organizations because of a security weakness showing that any company or federal agency associated with another is at risk when a hacking occurs).


One possible scenario is that hackers could get into U.S. government computers and then send false requests to the team. Another possible scenario is that an insider could be compromised and subsequently abuse their position. Abelson et al., supra note 6, at 2.

Id.
See BERKMAN CTR. FOR INTERNET & SOCIETY, HARV. UNIV., DON'T PANIC: MAKING PROGRESS ON THE “GOING DARK” DEBATE (2016), https://cyber.law.harvard.edu/pubrelease/dont-panic/Dont_Panic_Making_Progress_on_Going_Dark_Debate.pdf (questioning the “going dark” FBI approach by questioning whether the ability to surveil criminals and bad actors will be impossible and determining that would not be the case).

According to the authors of Don't Panic, factors includes the combination of technological developments and market forces (think: targeted advertising), and the “prevalence of network sensors and the “Internet of Things.” Id. at 19.

See id. at 2 (arguing that the market will cause an increase in unencrypted data, filling the encryption gaps that governmental agencies such as the FBI assume will go dark in the future).

The Ground Truth About Encryption and the Consequences of Extraordinary Access, THE CHERTOFF GROUP 2 https://chertoffgroup.com/cms-assets/documents/238024-282765.groundtruth.pdf (last visited Oct. 18, 2016) [hereinafter Ground Truth]. In the interest of completeness, Ground Truth continues on “... but in the future it is likely that ubiquitous encryption will have an impact on law enforcement capabilities.” Id. This is contradictory to the central thesis of Don't Panic that ubiquitous encryption is unlikely to occur due to certain market forces. Mr. Chertoff, the founder of the Chertoff Group, is the former Secretary of Homeland Security, during the George W. Bush Administration. He was also formerly a Federal Judge on the Third Circuit Court of Appeals. The Honorable Michael Chertoff, THE CHERTOFF GROUP https://chertoffgroup.com/bios/michael-chertoff.php (last visited Oct. 18, 2016).

General Hayden is the former Director of the NSA and the CIA. He was also the Principal Deputy Director of National Intelligence, amongst other prestigious posts in the Defense and Intelligence sectors. He is currently a Principal in the Chertoff Group. Michael Hayden, THE CHERTOFF GROUP https://chertoffgroup.com/bios/michael-hayden.php (last visited Oct. 18, 2016).

Ground Truth, supra note 266, at 17.

See id. at 17-18 (stating that the cables will cost $185 billion).

Id.

See, Cushing, supra note 221 (providing a pertinent example of the lack of a unified front on encryption between the Department of Defense, the NSA, and the FBI, three of the key players in security).
SHEDDING LIGHT ON THE “GOING DARK” PROBLEM AND THE ENCRYPTION DEBATE

ABSTRACT

In an effort to protect the enormous volume of sensitive and valuable data that travels across the Internet and is stored on personal devices, private companies have created encryption software to secure data from criminals, hackers, and terrorists who wish to steal it. The greatest benefit of encryption also creates the biggest problem: Encryption software has become so secure that often not even the government can bypass it. The “Going Dark” problem—a scenario in which the government has obtained the legal authority to search a suspected criminal’s encrypted device but lacks the technical ability to do so—is becoming increasingly common. In response, the government has resorted to obtaining court orders to compel private companies to assist it in bypassing encrypted devices, in some cases demanding that companies create entirely new software to accomplish this task. This raises a plethora of political, economic, and legal questions. This Note argues that given the weighty interests on all sides of the debate and the widespread effects that these cases will have, the encryption issue should be decided by the legislative branch instead of the courts. Because of the complexity of these issues and the lack of current legislation, the courts are being forced to stretch the law in ways that will likely lead to inconsistent and undesirable rulings. This Note advocates that the best method for Congress to solve this problem is to create an administrative body with rule-making, investigative, and adjudicative powers to address these situations on a case-by-case basis and to advise Congress on future legislation regarding encryption and digital security in general.

INTRODUCTION

People worldwide have become increasingly reliant on smartphones, tablets and personal computers, so much so that these devices have become an integral part of everyday life. Individuals store data including private conversations, photos, financial information, health data, passwords, places of residence, information regarding the whereabouts of themselves and others, and many other forms of extremely personal information on these devices. These hubs of personal information attract criminals and hackers who wish to use such information without the consent of the owner. In an effort to protect individuals’ personal information, private companies spend a considerable amount of time and resources encrypting it.
The greatest benefit of encryption also creates the biggest problem: private companies are very good at preventing third parties from bypassing their encryptions. While this can be advantageous for law-abiding citizens who need to be protected from hackers, it can be dangerous when criminals and terrorists are able to use the same encryption methods to communicate secretly, hide evidence, and otherwise evade authorities and commit crimes. Problems and conflicts arise when the government seeks to obtain this information for the purposes of investigating or thwarting crimes but cannot bypass the encryption of the device where the information is stored. Companies have become so skilled at creating device encryption that the government is often unable to access encrypted information even when it has the accompanying devices in its possession. 

Recently, the government has resorted to compelling the assistance of private companies to decrypt, or unlock, devices by obtaining court orders pursuant to the authority of the All Writs Act of 1789 (“AWA”). These companies have recently started resisting these orders on two grounds. First, private companies argue that the government’s interpretation of the AWA is incorrect because the statute does not actually grant the government the power to compel private actors to assist in decrypting devices. Second, companies argue that compelling them to do so is unconstitutional.

This Note analyzes the competing interests in the context of device encryption, describes the inadequacy of current law, and suggests a legislative solution to the existing conflicts and issues. Part I describes the basic workings of encryption, the “Going Dark” problem, and the tension between device encryption and the government’s interest in upholding national security. Part I also explores the recent court case between Apple, Inc. (“Apple”) and the Federal Bureau of Investigation (“FBI”) and describes its significance to this issue. Part II analyzes the current state of United States law regarding device encryption. Part II also examines relevant case law and the United States Constitution to determine the limits of the government’s power to compel the assistance of private companies to decrypt devices. Part III advocates that Congress pass new legislation, creating an administrative body equipped with rulemaking ability and adjudicative powers to determine on a case-by-case basis when private companies must assist the government in decrypting devices.

I. COMPETING INTERESTS: PRIVACY AND SECURITY IN REGARD TO ENCRYPTION

A. Background Information on Encryption and Its Importance

Encryption is commonly defined as “the conversion or encoding of information for transmission so as to prevent interpretation without the key for decryption.” Encryption enhances the security of a device by scrambling its contents so that only someone with the correct encryption key can decipher them. Encryption is generally regarded as the most effective way to protect confidential information. According to Dave Anderson, senior director of Voltage Security, an encryption technology provider, “[w]hen properly implemented, encryption provides essentially unbreakable security.” Device encryption is such a popular means of securing confidential information that technology companies constantly compete to have more secure products to respond to consumer demand. Some companies focus exclusively on digital security and encryption, such as companies like Symantec Corporation, while others, such as Apple, develop their own encryption technologies and use them in conjunction with their own products.

As technology advances, sophisticated hackers and cybercriminals are becoming increasingly effective at bypassing encryption and cyber intrusions are increasing in frequency. The cost of having weak encryption, or failing to encrypt at all, can be devastating. For example, in 2007, cybercriminals gained access to more than 100 million credit and debit card numbers from Heartland Payment Systems, Inc. In October 2014, attackers hacked Sony Pictures and released confidential information ranging from employee personal data to documents about the company’s upcoming projects.
While it was never confirmed, the FBI presented evidence that North Korea was behind this attack on Sony Pictures, but also noted, according to CNN, that “[a]nyone could have pulled this off .... It could have been a disgruntled Sony Employee, profit-seeking hackers, North Korea--or the combination of the three.”

Even more recently, the government itself was hacked when the Office of Personnel Management was hacked by criminals in 2015, putting millions of Americans' personal information at risk. The breach, which constitutes the largest cyberattack on the United States government in history, resulted in the theft of sensitive information from 21.5 million people, including addresses, health and financial history, fingerprints, and other information. When asked about the significance of the breach, James B. Comey, director of the FBI, stated, “[I]t is a treasure trove of information about everybody who has worked for, tried to work for or works for the United States government .... Just imagine you are an intelligence service and you had that data, how it would be useful to you.” Such information could be used or sold to those who seek to harm the victims of the theft or to infiltrate other governmental operations. These examples highlight the ease and frequency with which hackers are able to take advantage of weak encryption and cybersecurity systems.

Thus, both private companies and the government invest a substantial amount of time, money, and personnel to continuously make encryption systems stronger to stay one step ahead of hackers. Creating and updating encryption systems is no easy task; it requires that the software be designed, created, tested, validated, and deployed, which all take significant time and skilled technicians to accomplish. As a result of this investment, encryption has become so effective that it has caused an entirely new problem: criminals, hackers and terrorist can use the same encryption to their advantage.

**B. The “Going Dark” Problem**

The “Going Dark” problem, a phrase coined in the 1990s, refers to situations in which the government has legally obtained the right to search certain devices, but “lacks the technical ability to carry out those orders because of a fundamental shift in communications services and technologies.” According to the FBI, the Going Dark problem presents a huge obstacle for law enforcement trying to conduct investigations. FBI Director James Comey commented on the Going Dark problem by stating that “[a]rmed with lawful authority, we increasingly find ourselves simply unable to do that which the courts have authorized us to do, and that is to collect information being transmitted by terrorists, by criminals, by pedophiles, by bad people of all sorts.”

The FBI also claims that these problems are exacerbated in time-sensitive situations when decrypting a device could lead to thwarting or preventing a crime or saving lives. In order to overcome the Going Dark problem, the government has resorted to using the courts to obtain orders to compel private companies to assist them in their decryption efforts.

**C. Competing Interests**

The government's demand that private companies assist it in accessing encrypted devices has created extreme tension between the competing interests of those involved. These interests include those of the government and law enforcement, encryption companies, and the many individuals who rely so heavily on these devices. The relevant competing interests include national security in terms of public safety, the potential backdoor problem, privacy, national security in terms
of data security, and the constitutional concerns that will be covered in Part II. In order to find the proper solution to this problem, it is first necessary to consider the weighty interests on all sides of the debate.

1. National Security: Public Safety

In his testimony before the Senate Judiciary Subcommittee on the Constitution, Federalism and Property Rights, New York University School of Law professor Richard A. Epstein stated, “No one can doubt the legitimate needs of law enforcement officials at the federal, state and local levels to monitor the high tech criminal activities that threaten the security of this nation, the liberty of the *495 citizens within it, and the security and safety of the property they own.” 25

Criminals and terrorists alike use encrypted devices because of the difficulty the government has decrypting them. 26 For example, at the state level, according to Cyrus Vance, the Manhattan District Attorney, “[m]ore than 120 Manhattan criminal cases have been harmed by the failure to execute search warrants on the latest smartphones.” 27 Vance even noted that there were several occasions in which “average criminal [s]” informed their associates that only the latest encryption software (iOS 8 at the time), would prevent law enforcement from decrypting devices. 28 Furthermore, the Federal Law Enforcement Officers Association noted that in an intercepted phone call from a New York prison, an inmate referred to Apple iPhone encryption software as a “gift from god” because authorities were powerless to decrypt it. 29

This problem is not limited to domestic crimes. For example, as reported by the Washington Post, FBI Director James Comey stated, “[t]he Islamic State terrorist group is increasingly using encrypted communications to recruit troubled Americans and urge them to carry out attacks ...” 30 According to Comey, the Islamic State operatives post on social media outlets such as Twitter to attract potential recruits. 31 Once the operatives locate a potential recruit, they instruct him or her to use an encrypted mobile messaging app, at which point the FBI is no longer able to intercept messages or otherwise monitor their activity; all further communication between operatives and potential recruits becomes undetectable to  *496 the FBI. 32 The Washington Post noted that Comey stated to Congress that “the Islamic State has attracted at least 21,000 English-speaking followers on Twitter, bombarding them with incitements to violence.” 33 Thus, FBI officials advocate for the ability to compel private companies to provide reasonable assistance to help minimize these substantial national security risks. 34

2. Private Actors’ Data Security and “Backdoor” Concerns

Private individuals and corporations have a strong interest in data security. Individuals store a wealth of private information on their various electronic devices including, “financial records and credit card information, health information, location data, calendars, personal and political beliefs, family photographs, [and] information about their children” that if stolen by hackers could put them and their loved ones at risk. 35 According to a report from the Ponemon Institute, “[a]n estimated 47 percent of all American adults have been affected by data breaches over the last year [2013], with an estimated 432 million online accounts being affected.” 36 Companies are vulnerable to cybercrime as well; in fact, according to the FBI, “companies are the primary victims of cyber intrusions ....” 37 Indeed, in 2014, Target Corporation was the victim of one of the largest breaches of consumer data in American history when hackers gained access to over 40 million customer accounts, which included information such as credit and debit card numbers. 38
Since Target's data breach, major data breaches have been discovered almost every month, including breaches of several large companies including Michaels Stores, Sally Beauty Supply, Neiman Marcus, AOL, eBay, and P.F. Chang's Chinese Bistro. According to a report from Intel Security and the Center for Strategic and International Studies, “cybercrime is costing the global economy $575 billion and the U.S. economy $100 billion annually ... making the U.S. the hardest hit of any country.” The increased threat of cybercrime led companies such as Apple, Google, and Facebook to begin encrypting their devices and communication platforms by default beginning in 2014.

The government's request to compel companies to create new software to bypass its own encryption may lead to what many companies and cryptology experts are referring to as a “backdoor” or master key. Apple stated in a letter to its customers:

The government suggests this tool could only be used once, on one phone. But that's simply not true. Once created, the technique could be used over and over again, on any number of devices. In the physical world, it would be the equivalent of a master key, capable of opening hundreds of millions of locks.

Several leaders in the technology industry, including Amazon.com, Box, Cisco Systems, Dropbox, Evernote, Facebook, Google, Microsoft, Mozilla, Nest, Pinterest, Slack, Snapchat, Whatsapp, and Yahoo, filed a brief in which they collectively supported Apple's claim that the government's action would necessarily lead to the creation of a backdoor. According to a group of leading cryptographers and computer scientists, “[a]ny trap door system increases the risk that someone else will be able to find, duplicate or manufacture the key to the encrypted information.”

Furthermore, these companies and Apple argue that neither Apple nor the government could guarantee the safety of this software because creating the backdoor creates an extremely high risk that criminals, terrorists, or even foreign countries will obtain and use the backdoor to harm the government, citizens, or the companies themselves. In short, these companies claim that the risk in creating a backdoor is that once created, it is too difficult to secure it and to ensure that no one else comes through it.

The aforementioned technology leaders and Apple also claim that--aside from the security issues--creating a backdoor is an extremely burdensome and costly process.

Apple claims the government's unprecedented demand requires that it develop new software that will destroy the security features Apple has spent years creating. According to Apple, no current operating system can accomplish what the government desires, and any effort to create such an operating system requires that Apple write new code, rather than simply disabling the functionality of existing software. In its Motion to Vacate, Apple stated:

Experienced Apple engineers would have to design, create, test, and validate the compromised operating system, using a hyper-secure isolation room within which to do it, and then deploy and supervise its operation by the FBI to brute force crack the phone's passcode. The system itself would have to be tested on multiple devices to ensure that the operating system works and does not alter any data on the device. All aspects of the development and testing processes would need to be logged and recorded in case Apple's methodology is ever questioned.
Finally, Apple and other technology leaders have a strong interest in ensuring the security of hundreds of millions of customers who depend on them. The government forcing these companies to weaken their own encryptions will cause their customers to lose confidence in these companies' ability to protect user data from criminals, hackers, and terrorists, as well as the government.

3. National Security: Data Security

The government has a very strong interest in upholding national security, and that interest necessarily requires strong encryption. The FBI webpage states, “[t]he FBI supports strong encryption, and we know firsthand the damage that can be caused by vulnerable and insecure systems. The government uses strong encryption to secure its own electronic information, and it encourages the private sector and members of the public to do the same.” Secretary of Defense Ashton B. Carter noted at an annual computer security conference in San Francisco: “Data security, including encryption, is absolutely essential to the Pentagon.” Furthermore, Mr. Carter made it clear that he is “more interested in securing data than prying into it.” Like all private citizens and companies, the government is also at risk from hackers, criminals, and terrorists that seek to steal private information.

The government itself has had problems in the past with data security. As discussed earlier in this Note, the federal government--specifically the Office of Personnel Management--fell victim to a breach of government computer systems in 2015 that resulted in the theft of the personal information of approximately 21.5 million people. In the hacking, 19.7 million people who had undergone government background checks in the past and 1.8 million others, including spouses and friends, had their personal information stolen. In a separate but related hacking, 4.2 million federal employees also had their personal information stolen. These coordinated breaches constitute the largest known cyberattack on the systems of the United States government.

It is clear that the government has a strong interest in the preservation of national security by having both the means to access encrypted devices and the ability to use stronger encryption to secure its own sensitive information. Administrative officials stated, “The United States government firmly supports the development and robust adoption of strong encryption, which is a key tool to secure commerce and trade, safeguard private information, promote free expression and association .... At the same time, encryption poses a grave challenge for our national security and law enforcement professionals.” This challenge--the government's difficulty bypassing the encrypted devices of suspected criminals, hackers, and terrorists--is fueled by the fact that whenever advancements in device encryption can be used to better protect public and private users, criminals, hackers, and terrorists can usually also use the same technology for more unsavory purposes.

Some government officials claim that both interests can be accomplished and are not mutually exclusive. Other officials claim that allowing the government to more easily bypass encrypted devices necessarily weakens encryption and that promoting stronger encryption necessarily exposes the government to the aforementioned national security risks; in other words, that both interests cannot be reconciled. The Justice Department and the FBI are advocating for a compromise in the form of promoting strong encryption “with limits” to advance both interests. The difficulty lies in determining what the limits should be and whether such a compromise can truly accomplish both interests.
C. Case Study: San Bernardino Case

The issue of device encryption has resurfaced in the form of the FBI demanding, via a court order, that Apple, a private company, assist in its investigation of a terrorist attack that took place in San Bernardino, California. In In re the Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, 66 (“San Bernardino case”), the FBI persuaded the court to grant a Motion to Compel Apple to reasonably assist in the FBI's investigation by decrypting Syed Rizwan Farook's Apple iPhone to obtain personal information that might relate to his participation in the terrorist incident. Though in the past Apple has complied with such orders, it recently decided to appeal the court's ruling, claiming that “[t]he U.S. government has asked us [Apple] for something we simply do not have, and something we consider too dangerous to create. They have asked us to build a backdoor to the iPhone.” 67 According to Apple, the government's actions constitute an overreach of its authority and an attempt to subvert the law-making process by using the courts as a means to authorize its actions as opposed to using the normal lawmaking process. 68 Apple further contends that being forced to create a backdoor would not only put its users at risk, but would undermine the very encryption systems it has gone to such great lengths to create.

1. The Significance of the San Bernardino Case

The San Bernardino case is significant because it is the first case in which the government attempted to bypass encryption by demanding that a private company create entirely new software in order to do so. In its attempt to heighten encryption security, Apple released a new operating system called iOS 8, which effectively prevented Apple itself from being able to decrypt devices running iOS software. 69 Once any Apple device is running iOS 8 or a later update, not even Apple can access the information stored on the locked device as opposed to before the update, when Apple had the ability to decrypt devices running iOS software. 70 In its letter to its consumers, Apple stated, “We have even put that data out of our own reach, because we believe the contents of your iPhone are none of our business.” 71

The iPhone in the San Bernardino case, unlike the other cases, is running iOS 8, and therefore the only way Apple or the government would be able to access the device is by creating an entirely new program to decrypt it, commonly referred to as a backdoor. 72 In other words, whereas in the past Apple had access to the information and could simply hand it over to the government, it is now being asked to create a new program that would subvert its own security measures and thereby give the government the information it seeks. As this Note will explain, there are significantly different legal implications when the government requests that a private company create entirely new software as opposed to when it simply requests information the private company already has in its possession.

II. CURRENT STATE OF THE LAW

This Part addresses the sources of law that might authorize the government to compel private actors to assist it in bypassing device encryption. This Part also addresses whether the applicable statutes are being applied constitutionally and other relevant constitutional concerns. This Part concludes by commenting on the difficulty *503 courts have applying the limited and antiquated statutes and case law surrounding this issue.

A. What Source of Law Might Authorize the Government to Compel Private Companies to Provide Assistance to Bypass Device Encryption?
The two primary statutes at issue are the AWA and the Communications Assistance for Law Enforcement Act ("CALEA"). In the several cases in which the government has demanded the assistance of private companies to bypass device encryption, it has relied on the authority of the AWA. The Supreme Court has noted that “where a statute specifically addresses the particular issue at hand, it is that authority, not the All Writs Act, that is controlling.” The AWA is considered to be a “gap-filler” and was created to endow the courts with “broad statutory authority to ensure they could effectively carry out the duties of an independent judiciary by issuing the orders necessary to do so—even if Congress had not had the foresight to create all of the procedural mechanisms that might be required.”

Apple argues that CALEA applies to the matter at hand and is therefore controlling instead of the AWA. Conversely, the government argues that the AWA is controlling. The issue is whether CALEA addresses the specific matter at hand, or if instead the AWA is controlling because neither CALEA nor any other statute directly addresses this matter. Because the AWA only applies in the absence of any other controlling statute, it is first necessary to determine whether or not CALEA applies in this situation.

1. Communications Assistance for Law Enforcement Act (CALEA)

In 1994, Congress passed CALEA to address its concern that “new and emerging telecommunications technologies pose problems for law enforcement.” CALEA has several limitations regarding what the government can and cannot compel third parties to do when providing technological assistance. In short, “CALEA requires telecommunications carriers to retain the ability to comply with court orders for real-time interceptions and call-identifying information.”

In the San Bernardino Case, Apple argued that CALEA explicitly prohibits the government from compelling companies like Apple to decrypt devices and, therefore, that the government is precluded from using the AWA to circumvent CALEA. CALEA states that the government cannot compel a provider of an “electronic communication service” to adopt any specific design of its equipment, facilities, services or system configuration. Apple argues that it is “unquestionably a provider of electronic communications services through the various messaging services it provides to its customers through iPhones.” Thus, Apple argues the government cannot require it to create new software, as that would constitute adopting a specific design of its equipment, facilities, services or system configuration. According to Apple, because CALEA specifically prohibits this, the government cannot use the AWA to circumvent CALEA.

Apple further argued that applying the AWA here would be “inconsistent with the intent of Congress.” Apple noted that CALEA contains mandatory assistance provisions, which list the types of actors that must assist the government. Apple stated, “CALEA intentionally excludes information services providers like Apple, from the scope of its mandatory assistance provisions.” Apple contended that because Apple does not fall into one of the categories of actors listed, Congress did not intend for the government to be able to compel companies like Apple to assist it. Indeed, in another similar case involving the FBI compelling Apple to assist it, the district court noted, “CALEA ... is part of a larger legislative scheme that is so comprehensive as to imply a prohibition against imposing requirements on private entities such as Apple that the *statute does not affirmatively prescribe.*” Apple therefore argues that even if CALEA does not explicitly prohibit the government from compelling Apple to assist it, because Congress has indicated its intent to “allow strong encryption” and to prevent the government from “mandating that such encryption schemes contain a back door,” the courts should refrain from extending CALEA to apply in this case.
The government argued that CALEA does not forbid the court's order compelling Apple to provide reasonable assistance for several reasons. First, the government argued that CALEA does not “meticulously, intricately, or specifically” address when a court may require a private company such as Apple to help the government bypass encryption software, and therefore did not specifically address the matter at hand. The government argued that the AWA applies as long as there is no statute that explicitly controls, which is the case here. The government stated, “It is not enough for other laws to brush up against similar issues. Rather, Congress must legislate so ‘intricately’ as to leave ‘no gap to fill.'” Because of this silence or gap in the law, the government claimed that an order under the AWA was appropriate to compel Apple to provide assistance. The government also argued that CALEA only limits the authority of law enforcement agencies and not courts. Indeed, the government argued, “The Order rests not on CALEA, but on the AWA ....” Finally, the government noted that CALEA only applies to data “in motion.” Data in motion is a phrase used for data that is in transit, especially via telephone or the Internet. Generally, data that is still being transmitted is considered in motion, whereas data that is stored on a device and no longer being transmitted is consider at rest. According to the government, because it sought data that was stored on a device and no longer being transferred (data at rest), the plain text of CALEA does not apply. The government argued that whether Apple is an information service provider as defined by the statute is irrelevant because the statue only applies to data in motion. The government stated, “Put simply, CALEA is entirely inapplicable to the present dispute and does not limit this Court's authority under the All Writs Act to require Apple to assist the government in executing a search warrant.”

It is unclear whether or not CALEA directly applies to the matter at hand. According to Magistrate Judge Orenstein--a United States Magistrate Judge for the Eastern District of New York who ruled on another similar matter involving the FBI and Apple--it is arguable whether or not CALEA explicitly prohibits the government from compelling companies like Apple to assist it. It does at least seem to be clear that because the data in question is at rest rather than in motion, the plain text of CALEA does not explicitly prohibit the government from compelling Apple to assist it in its investigation with a court order pursuant to the AWA.

2. The All Writs Act of 1789

In the several cases in which the government has demanded the assistance of private companies to bypass device encryption, it has primarily relied on the authority of the AWA of 1789. The AWA states, “The Supreme Court and all courts established by Act of Congress may issue all writs necessary or appropriate in aid of their respective jurisdictions and agreeable to the usages and principles of law.” The statute grants federal courts the authority to issue orders when three criteria are met:

1. Issuance of the writ must be “in aid of” the issuing court's jurisdiction;

2. The type of writ requested must be “necessary or appropriate” to provide such aid to the issuing court's jurisdiction; and

3. The issuance of the writ must be “agreeable to the usages and principles of law.”
If an application under the AWA meets all three of those requirements, the court “may” issue the requested writ in exercise of its discretion--but it is never required to do so. 102

After a court has determined that the statutory requirements of the AWA have been met, it may then consider the discretionary factors laid out by the Supreme Court in United States v. New York Telephone Company. 103 In N. Y. Tel. Co., the Court ruled that a private telephone company could be required to install pen registers (surveillance devices that record outgoing phone numbers when dialed) in order to assist the government in thwarting criminal activity. 104 In that case, the Court ruled that three discretionary factors should be taken into account when considering an application pursuant to the AWA: (1) “the closeness of the relationship between the person or entity to whom the proposed writ is directed and the matter over which the court has jurisdiction”; (2) “the reasonableness of the burden to be imposed on the writ's subject”; and (3) “the necessity of the requested writ to aid the court's jurisdiction (which does not replicate the second statutory element, despite the overlapping language).” 105 The requirements of the AWA itself seem fairly straightforward, but the constitutionality of the government's use of the AWA still must be scrutinized.

B. Constitutionality of Deployment of the AWA

The courts have made several rulings regarding the constitutionality of the AWA. In Pennsylvania Bureau of Correction v. United States Marshals Service, the Supreme Court ruled that “[t]he All Writs Act is a residual source of authority to issue writs that are not otherwise covered by statute. Where a statute specifically addresses the particular issue at hand, it is that authority, and not the All Writs Act, that is controlling.” 106 Earlier, in Plum Creek Lumber Company v. Hutton, the Ninth Circuit held that “[t]he All Writs Act is not a grant of plenary power to the federal courts .... It does not authorize a court to order a party to bear risks not otherwise demanded by law, *508 or to aid the government in conducting a more efficient investigation, when other forms are available.” 107

Finally, according to Magistrate Judge Orenstein, the AWA may be appropriately used “to fill in a statutory gap that Congress has failed to consider, [but] it cannot be used to grant the government authority Congress chose not to confer.” Judge Orenstein explained as follows:

But if CALEA, considered in the context of a larger statutory scheme, does not erect such a barrier to relief on its own terms, then the Application turns on whether the gap in the laws the AWA fills is, as the government argues, the entire space between authorizing statutes and legislative prohibitions or if, as Apple would have it, it only reaches to such legislative powers as Congress has not considered and either adopted or rejected. 108

The question is whether Congress's inaction reflects a purposeful choice not to authorize the courts to compel information service providers to provide assistance to the government or if it is simply congressional oversight.

I. First Amendment Implications

One constitutional concern this use of the AWA raises is freedom of expression via software code. It might be the case that by forcing companies to write specific software code, the government is compelling speech and therefore violating
the First Amendment of the Constitution. In *Turner Broadcasting Systems Inc. v. Federal Communications Commission*, the Supreme Court ruled that government actions that “compel speakers to utter or distribute speech bearing a particular message are subject to the same rigorous scrutiny” as laws restricting speech. In *Riley v. National Federation of the Blind of North Carolina, Inc.*, the Supreme Court held that compelled speech is a “content-based regulation of speech” subject to strict scrutiny. In *Junger v. Daley*, the Sixth Circuit ruled that “[b]ecause computer source code is an expressive means for the exchange of information and ideas about computer programming, we hold that it is protected by the First Amendment.” Finally, in *Universal City Studios, Inc. v. Corley*, the Second Circuit determined that computer code is treated as speech with regard to the First Amendment.

Apple and the technology industry leaders argue that these cases applied together require the courts to view the government's action under a strict scrutiny basis. They argue that the government is regulating code, which is content-based speech, and that by compelling private companies to write specific code, the government is violating those companies' First Amendment rights.

The government, on the other hand, contends that the code it is requesting is functional in nature rather than expressive and therefore does not receive the full protection of the First Amendment. The government relies on the Court's holding in *Universal City Studios v. Corley* that “solely functional code 'is not within the meaning of the First Amendment.'” The government also claims that based on the holding in *Red Lion Broadcasting Company v. Federal Communications Commission*, 395 U.S. 367, 386 (1969), “software that is automatic and is to be used in an entirely mechanical way is not speech under the First Amendment.”

In the San Bernardino case, the government argues that based on the aforementioned cases, the Court should analyze motions to compel companies to create new software under an intermediate scrutiny test. Therefore “so long as the regulation services a substantial government interest[,] the interest is unrelated to the suppression of free expression and any incidental restrictions on speech must not burden substantially more speech than is necessary to further that interest.” The government asserts that its interest in obtaining the code is to thwart terrorist threats, which is both substantial and does not relate to the suppression of speech, and that it does not incidentally burden any speech unnecessarily because there is no other way the government can bypass the companies' encryptions.

### 2. Fifth Amendment Implications

Another constitutional concern raised by this use of the AWA is the Fifth Amendment's guarantee of substantive due process. In the San Bernardino case, Apple argued the government violated its substantive due process right to be free from “arbitrary deprivation of its liberty by government.” Apple relied on the ruling in *County of Sacramento v. Lewis*, in which the Court stated, “We have emphasized time and again that '[t]he touchstone of due process is protection of the individual against arbitrary action of government,' ... [including] the exercise of power without any reasonable justification in the service of a legitimate governmental objective.”

The government responded that “Apple must also show that such deprivation was ‘clearly arbitrary and unreasonable, having no substantial relation to the public health, safety, morals, or general welfare.’” The government argued that this cannot be the case because its actions have a substantial relation to public safety and general welfare since it is using the information to conduct an FBI investigation aimed at thwarting future terrorist attempts and apprehending any accomplices to the attack.
3. There is No Fourth Amendment Argument

At first glance, the government decrypting people's personal devices may seem to raise Fourth Amendment issues, but that is not the case. Although encryption is created to address its users' personal privacy concerns, the Fourth Amendment's protection against an unlawful search and seizure does not apply here. Indeed, when discussing the San Bernardino case, City University of New York (CUNY) School of Law Professor Ruthann Robson briefly addressed the issue stating that, “Interestingly, there is no Fourth Amendment argument.”

The Fourth Amendment prohibits unreasonable searches and seizures and requires a judicially sanctioned warrant supported by probable cause before a search or seizure can be conducted. The AWA requires a search warrant be obtained before a court can grant any motion pursuant to the statute. Because the government must first obtain a search warrant before using the authority of the AWA to obtain a motion to compel private companies to assist it, the Fourth Amendment requirements have been satisfied. According to a Congressional Research Report, “Since the government obtained a valid probable cause warrant in this case, Apple is not contesting the search of the device under the Fourth Amendment. Thus, any privacy interest involved must derive from some other constitutional, statutory, or extra-constitutional source.” The matter at hand poses little threat to individual privacy rights because, in order for the AWA to be utilized, the government must first obtain a search warrant and therefore establish probable cause, which satisfies the Fourth Amendment. There are, of course, still privacy concerns in the sense that individuals have an interest in having their data protected, but there is no Fourth Amendment claim.

As exemplified by these arguments, courts struggle to determine what statutes to apply, how to apply them, and whether such applications violate the constitution. The current statutory framework provides less than clear guidance on balancing the needs of companies and the government in an area of continual technological advancements and pervasive tension between security and privacy. This problem is exacerbated by the fact that these cases are generally very fact-specific and technical, making it difficult for courts to apply the law consistently and in a way that takes into account the various interests involved.

*512 III. REFORMS AND SOLUTIONS

The conflicting interests of the need for increasingly stronger encrypted devices and of the government to thwart cybercrimes cannot be reconciled within the current system. The Apple case highlights how complex these issues can be, as well as how difficult it can be for a court to determine the proper outcome of these cases. Without reform, the Going Dark problem and cybercrime will continue to plague private actors and the government alike. This Part will first describe why the judicial branch is not the proper branch to solve issues of this kind. This Part will then propose that Congress should create an agency capable of addressing these concerns more nimbly than Congress and with more accuracy than the courts. Finally, this Part will address potential criticisms of the proposed law reform.

A. The Courts Alone Are Not the Proper Branch for Determining the Outcomes of These Situations

1. Congress is Better Equipped to Handle these Issues than the Judicial Branch

The judicial branch is not the branch best equipped to determine the outcome of compelled decryption cases. Rather, Congress should use its power to create an administrative body uniquely suited to handling Digital Security issues. There
are three primary reasons the courts are not well equipped to determine the outcome of these issues: (1) the issue is one of public policy; (2) the issue involves extremely complicated and ever-changing technology; and (3) the courts are limited to applying antiquated or ill-suited law.

First, to arrive at a sensible result, the decision maker must, in each case, make a holistic ruling by taking into account issues such as national security, public safety, data security, individual privacy, innovation, and corporate competitiveness. The courts are not the proper venue to balance these interests, as they are limited to considering the interests of parties involved and the relevant law at the time. This decision will affect every person, corporation, and government entity that uses encryption software. It is well-established that the legislative branch is the preferable governmental branch to determine issues that primarily concern public policy. Unlike the judicial branch, the legislative branch can utilize tools such as appointing committees, inviting and considering public discourse and lobbying, amending statutes, and--as this Note advocates--delegating responsibilities to administrative bodies. Indeed, according to one constitutional and international law professor at Harvard Law School, Noah Feldman, “courts that are pretty good at interpreting statutes or applying the Constitution generally aren't very good at identifying and weighing major domestic and international public policy consequences .... In the U.S. system, Congress is supposed to make difficult public policy decisions.” Furthermore, the Supreme Court in INS v. Chadha noted that Congress is uniquely suited to make laws because of the Framers' decision “that the legislative power of the Federal Government be exercised in accord with a single, finely wrought and exhaustively considered procedure.”

Second, the technological aspects of these matters are extremely complicated and rapidly evolving, such that the courts will not be able to keep pace. Indeed, several leaders in the technology industry collectively stated that “[i]n light of rapidly evolving technology and its tremendous social benefits, Congress is better suited to confront the issues here.” Congress has the power to call upon experts, create committees, hear from various lobbyists, and address situations without needing to wait for a party to file a claim. Unlike the courts, Congress is able to take into account factors outside of the case at hand and weigh the benefits and harms to the public at large, whereas the courts are limited to only considering the parties directly involved and the applicable statutes.

The aforementioned Eastern District of New York Apple case demonstrates that the legislative branch is better equipped than a court to determine the outcomes of these issues. In that case, Judge Orenstein's ruling was that none of the factors--the closeness of Apple to the matter, the burdensomeness on Apple, or the necessity of Apple's assistance--were such that it justified imposing the obligation to assist the government's investigation on Apple against its will. The problem is that none of the considerations in that analysis concern legal matters; they are all matters of fact, many of which deal with complex technological issues. For example, the burdensomeness factor would require the court to take into account all of the aspects involved with creating new software such as the cost of paying engineers to do so, the difficulty of accomplishing such a task, the equipment needed to complete it and the difficulty of acquiring said equipment, whether such a task would require moving to or even building a new facility, and many other considerations. As is evidenced by Judge Orenstein's ruling, technological facts are often the determining factors in these cases. Such factors can be extremely complicated, especially when dealing with software code, encryption software, backdoors, and other technological issues. Given this complexity, experts are increasingly needed to determine the correct results of these cases. Congress is better equipped to deal with this issue because it can--and should--create an administrative body with experts in the field to more accurately determine the correct outcome of these cases. Furthermore, because of the ever-changing landscape of technology, courts will always have the problem of being limited by legislation that will continue to become outdated. In a brief supporting Apple in the San Bernardino case, several security and cryptography experts stated, “The AWA's authority to issue writs to non-parties simply does not account for the public-security dangers this Court's Order creates, nor the future risks that future orders will also pose. The plain language of the statute creates
no obligations and gives no guidance to courts considering the very important and technologically nuanced underlying security risks associated with mandating forensic access to private data.”

While courts are limited to applying current legislation, Congress is the sole branch authorized to “update a technologically antiquated statute to address the new and rapidly evolving era of computer and cloud-stored, processed and produced data.”

It is worth noting that Congress is, however, also limited because each time it passes legislation it will quickly become outdated. Thus, the need for the creation of an administrative body is clear. Digital security technology grows at such a rate that the only way to effectively regulate it is to create a living, breathing administrative body that can make rules and decisions quickly and accurately enough to keep up with it.

Third, there is a clear lack of legislation, and, thus, the courts do not have sufficient tools to correctly determine the outcome of cases. For example, in the San Bernardino case, the only applicable statutes were CALEA, which may or may not apply, and the AWA, which is supposed to be used only as a gap-filler. It seems inconceivable that courts should be forced to determine issues pertaining to the latest encryption software protocols of 2016 and beyond armed with only a gap-filling statute written in 1789. Furthermore, according to Professor Feldman, “As written, the laws governing the granting of warrants don't provide sufficient latitude for a court to weigh the dangers of requiring [private companies such as] Apple to write new code against the corresponding gains for policing and national security.”

Finally, beyond the fact that Congress is better equipped to solve this issue, relying on the courts to do so violates the separation of powers doctrine.

2. Allowing the Courts to Determine this Matter Forces them to Violate the Separation of Powers Doctrine

It is well established that courts may not exercise the legislative power by repurposing statutes to meet the evolving needs of society. Allowing the courts to determine the outcomes of these cases, given the lack of legislation and strong public policy consequences, puts them in a situation in which they are responsible for weighing the aforementioned various interests and essentially determining an extremely important issue of public policy. The courts are being put in a position where they are essentially functioning as the legislature, which constitutes a violation of the separation of powers doctrine.

Indeed, in Baker v. Carr, the Supreme Court explained that courts should not make rulings that are, at bottom, “political questions.” The Court enumerated several factors to use in determining whether a question is a political question, including whether there is an “impossibility of deciding [the issue] without an initial policy determination of a kind clearly for nonjudicial discretion.” Courts determining this issue would have to decide if and when the government can compel private companies to assist it in investigations, which should constitute a “initial policy determination” as defined in Baker. Such policy determinations are better left to Congress.

The need for new clarifying legislation is clear. The reason there is so much debate about this issue is because Congress has not made a clear legislative decision, hence the argument about whether the AWA may be used to fill the missing gaps in these situations. Thus, the courts are put in a position where they must attempt to make the law by making rulings with limited statutory guidance and to potentially give new meaning to statutes instead of allowing Congress to update antiquated legislation or adopt new legislation. Indeed, Apple convincingly argued, “[Avoiding public debate] seems fundamentally inconsistent with the proposition that such important policy issues should be determined in the first instance by the legislative branch after public debate--as opposed to having them decided by the judiciary in sealed, ex parte proceedings.”
Finally, in *INS v. Chadha*, the Supreme Court held that “Congress has plenary authority in all cases in which it has substantive legislative jurisdiction, so long as the exercise of that authority does not offend some other constitutional restriction.” The fact that Congress has already grappled issues of encryption and digital security with legislation such as CALEA demonstrates that Congress has the authority and indeed the duty to solve this issue.

*517 B. Solution*

In order to correctly balance the many and heavily-weighted competing interests revolving around the issue of device encryption, Congress should enact a statute that creates an administrative agency with both rulemaking and adjudication powers pertaining to digital security, including encryption. An administrative adjudicative agency could solve these problems because it would not be constrained by many of the problems the courts and Congress face. In February 2016, several members of Congress introduced the Digital Security Commission Act of 2016 for the purpose of creating a temporary national encryption panel. The bill has only been introduced in the House of Representatives and, according to LexisNexis Bill Tracking, the bill has a “low chance to pass to the next stage.” Even if this bill were to pass, the Agency proposed in this Note (referred to as the “Digital Security Agency”) is unique in that, although it would seek to accomplish many of the same goals as the Digital Security Commission, it would go further than this bill does and would accomplish its goals by granting the agency power to make decisions. The Digital Security Commission Act of 2016 contains several useful ideas, however, it does not provide the national encryption panel any power to make decisions, and instead only serves to make periodic reports to Congress to advise it about the various benefits and risks associated with encryption technology. The Digital Security Agency would seek to accomplish many of the goals enumerated in Digital Security Commission Act of 2016, but instead of creating a panel, it would create an adjudicative agency that could hold hearings and make rulings rather than simply advise Congress. The statute this Note proposes would go much further than the Digital Security Commission Act by granting the agency the power to not only consult with experts in the field to make rules, but also to make adjudicative decisions to react to these situations more nimbly than Congress and with more resources than the courts. Finally, the Digital Security Agency would also be responsible for making periodic recommendations to Congress so that, in the event new legislation is needed, Congress will be better informed to respond.

*518 The Digital Security Agency would accomplish the purpose enumerated in the Digital Security Commission Act, but would have additional powers to accomplish these goals. The mission of the Digital Security Commission Act of 2016 is as follows:

To bring together leading experts and practitioners from the technology sector, cryptography, law enforcement, intelligence, the privacy and civil liberties community, global commerce and economics, and the national security community to examine the intersection of security and digital security and communications technology in a systematic, holistic way, and determine the implications for national security, public safety, data security, privacy, innovation, and American competitiveness in the global marketplace.

The Digital Security Agency would have two ways to make policy and two corresponding branches: an adjudication branch and a rulemaking branch. The adjudication branch would be responsible for holding adjudicatory hearings and making orders when the government seeks to compel private actors to assist it in its investigations. The rulemaking branch would be responsible for conducting research and consulting experts, private companies and the government in order to make rules and recommendations to Congress. By utilizing the two separate branches, the Agency would be more nimble than Congress, and more accurate than the courts.
1. The Adjudication Branch

The Digital Security Agency would have the power to hold hearings to determine on a case-by-case basis whether or not the government could—after obtaining a search warrant issued by a court—proceed to compel a third party to assist in decryption efforts. Essentially, law enforcement would have to first obtain a normal search warrant by showing probable cause. After having obtained a search warrant, if law enforcement found it was unable to decrypt a device, network, or other digital source, it would file a claim to the Digital Security Agency that would in turn adjudicate the matter between law enforcement and the private company. In other words, the Digital Security Agency's primary responsibility would be to respond when law enforcement encounters the Going Dark problem.

*519 The adjudication branch should be comprised of an odd number of qualified administrative law judges. In addition to the administrative law judges, there should be a council of experts to advise the judges on matters of fact. The council of experts should have members from the following areas: “(1) cryptography (2) global commerce and economics (3) federal law enforcement (4) state and local law enforcement (5) consumer-facing technology sector (6) enterprise technology sector (7) the intelligence community (8) the privacy and civil liberties community.” There should also be at least one member nominated by an interest group in support of private companies such as the Software Alliance or the Information Technology Industry Council. The council of experts would have no power to make the final rulings, however the administrative law judges should be advised to strongly consider the opinions of these experts when determining the outcomes of these matters. If the adjudicative branch determines by a simple majority that the government is entitled to compel a private company to provide assistance, it would issue an order stating that the company has to comply. The order would have the same effect as if a normal court granted a motion to compel a company to provide assistance. Companies could appeal this order to a federal court of appeals, but only on the grounds of a constitutional violation.

The enabling act should advise the administrative law judges to consider the factors laid out by Judge Orenstein: the closeness of the private company's relationship to the underlying criminal conduct and government investigation; the burden the requested order would impose on the private company; and the necessity of imposing such a burden on the private company. In addition, the enabling act should advise the adjudication branch to consider the probability that decrypting the device in question would lead to thwarting a crime or solving an investigation. The enabling act should, however, give the agency discretion to consider other factors, as new factors will almost certainly arise as technology advances. The orders the Digital Security Agency makes should be binding on the Agency itself so that it has to follow precedent, unless the rulemaking branch creates a rule that effectively overrules a decision. In the event new discoveries are made that have a substantial effect on a ruling the court has already made, the rulemaking branch would be responsible for creating a new rule to address the concern. This way, the adjudicative branch will ensure predictable results, but will also have the ability to quickly adapt to the ever-changing landscape of digital security.

The adjudicative branch is necessary to the agency because, while the rulemaking branch can conduct research and use notice and comment rulemaking, there must still be a method for the government to obtain an order compelling a private company to assist it. Regardless of how comprehensive the rules are, there will almost certainly be some disputes that must be resolved by an adjudicator. Furthermore, because technology changes so often, there will almost certainly be disputes for which there is no applicable rule. The adjudication branch will also be instructive for the rulemaking branch, which can compile adjudicative rulings into comprehensive rules. The adjudicative branch is essential to the agency's
ability to respond to disputes as they arise, and the panel of experts and the preference given to administrative law judges with some relevant expertise will lead to more accurate decisions than the courts could make.

2. The Rulemaking Branch

The primary responsibilities of the rulemaking branch would be to promulgate rules under the APA via notice and comment rulemaking pursuant to APA §553, and to provide periodic reports to Congress. This branch would be responsible for providing notice, the opportunity to comment, and publication of all rules with a general statement of basis and purpose of all rules adopted as proscribed by APA §553. By using notice and comment rulemaking, the agency will be able to more accurately ascertain the interests and concerns of those who will be affected by the rules. Rules made by this branch should be used by the adjudication branch in its rulings. This branch would also be responsible for submitting semi-annual reports to Congress detailing: (1) all agency action since the preceding report; (2) issues or controversies the agency believes present a problem or will become problematic in the foreseeable future; (3) an assessment of the most immediate and dangerous cyber security threats and potential methods to solve them; and (4) any information not specifically stated the agency believes is relevant and important.

This branch should be comprised of at least nine members for the purposes of making rules and sending reports to Congress regarding digital security and encryption. Like the panel of experts in the adjudicative branch, the rulemaking branch should be required to have at least one member with expertise in each of the following areas: “(1) Cryptography (2) Global commerce and economics (3) Federal law enforcement (4) State and local law enforcement (5) Consumer-facing technology sector (6) Enterprise technology sector (7) The intelligence community (8) The privacy and civil liberties community,” as well as a member with some federal lawmaking experience. There should also be at least one member nominated by an interest group in support of private companies such as the Software Alliance or the Information Technology Industry Council. This highly qualified team would be more informed than the average member of Congress, and could create rules more quickly because it would not be burdened by the procedures Congress must follow to pass bills.

The rulemaking branch provides the ideal mechanism for quickly making rules while considering all of the competing interests. First, the rulemaking branch will almost certainly lead to better rules than Congress alone could create. By utilizing notice and comment rulemaking, the branch will allow for greater public input into the decision, thereby ensuring the agency is adequately informed of all competing interests. This branch also will allow for more comprehensive decision making than the adjudicative branch. The rulemaking branch will also benefit the agency because it will be able to impose comprehensive rules at once on all similarly situated parties, as opposed to the adjudicative branch, which will be limited rulings on the parties involved in hearings. The qualified officers of the rulemaking branch will able to foresee problems before claims arise in some cases and will therefore be able to create rules the adjudicative branch can apply in the event that such a claim is eventually brought; this will decrease the likelihood that the adjudicative branch will have to “stretch” a rule to apply to a situation or give a rule new meaning, and will lead to more predictive outcomes. Finally, because rules must be published as soon as they are promulgated, regulated parties will have advance knowledge of their legal duties and can design their digital security protocols accordingly. Thus, the rulemaking branch is essential to the efficiency and comprehensiveness of the agency as a whole.

3. Appointment and Removal

The enabling act should specify that members for both the adjudicative branch and the rulemaking branch should be appointed by the President with the advice and consent of the Senate, pursuant to the Appointments Clause. In
order to be selected to serve as an administrative law judge, candidates should be required to meet all the qualifications enumerated by the APA. Special preference should be given to candidates who also have expertise in one of the aforementioned relevant areas of expertise. The enabling act should also restrict the removal of administrative officials and administrative law judges to “good cause” to insulate the agency from political pressure as much as possible. Finally, in the event of a vacancy in any position, that member should be replaced in the same manner in which he or she was appointed.

C. Potential Criticisms of the Agency Approach

Critics of the agency approach will likely argue that the courts are in fact better equipped than an administrative agency because courts are better at upholding constitutional rights, which constitutes an equal if not greater concern than the potential public policy consequences. In *Marbury v. Madison*, the Court established that all cases arising under the Constitution are to be reviewed by the judicial branch. The Court also stated, “It is emphatically the province and duty of the judicial department to say what the law is.” Critics may also argue that, oftentimes, agencies do not live up to expectations and fall short of the large responsibilities they are given. Particularly because of the high volume of anticipated cases and issues in the future, critics will likely be skeptical that this agency will be able to handle all of its responsibilities.

Admittedly, there are constitutional concerns and questions of law involved in some of these cases, but the proposed agency approach takes this into account by recommending that the agency include administrative law judges. The Administrative Procedure Act of 1946 was created to establish the administrative law judge function to “[e]nsure fairness in administrative proceedings before Federal Government Agencies.” The inclusion of administrative law judges would help to ensure that the agency correctly and fairly interprets issues of law. Furthermore, the inclusion of people with federal law making experience in the rulemaking branch will decrease the likelihood that any unlawful rules are created. Given that administrative law judges are as qualified as ordinary judges, there is no reason to believe the agency will not correctly and fairly determine laws and avoid making rulings contrary to the letter and spirit of the Constitution. Furthermore, there are plenty of other agencies already in place that make rules and regulations that relate to constitutional and legal concerns and are highly effective. There is no reason to believe that this agency would be less effective than the many successful agencies that are essential to the United States government.

Furthermore, all of the rules and decisions the agency makes can be reviewed by courts or overruled by a statute passed by Congress. Under the proposed agency approach, in the event that a party believes its constitutional rights have been violated, that party can always appeal the agency's rulings. In such a situation, the proposed agency would have still served its purpose, as the only appeals would be decisions in which there is a potential constitutional violation and not those revolving around the public policy concerns of encryption.

The ability of agencies with adjudicative judges to succeed where the courts have failed has been well documented, as exemplified by agencies such as the NLRB, which has been successfully engaging in adjudicatory and rulemaking activities for over 80 years. While the agency approach may be imperfect, given the nature and the difficulty of the issue, it seems to be a far superior approach than forcing the courts to determine the outcome of these matters with little to no legislative guidance.

CONCLUSION
Many believe that the question of reconciling competing encryption and digital security interests is one without an easy answer, and they are correct. The complexity of the situation and the interests at stake provide all the more reason for Congress to create an administrative agency to conduct trials as new issues arise and advise Congress on potential legislation for the future. If correctly implemented, the agency approach will ultimately provide the nation with the best policy for moving forward while simultaneously protecting the interests of national and digital security, privacy and public safety.

Footnotes

a1 J.D. Candidate, University of Michigan Law School, May 2017; B.A., Emory University, 2014. I would like to thank the editors of the Michigan Journal of Law Reform, Professor Leonard Niehoff, and my friend Stephanie Balitzer for their invaluable feedback.


2 See generally Order Compelling Apple Inc. to Assist Agents in Search, In re Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. 5:16-cm-00010-SP (C.D. Cal. Mar. 3, 2016) (exemplifying a circumstance in which the government cannot access an encrypted device despite having acquired a search warrant and the encrypted device).


7 Id.

8 See generally Brief of Amici Curiae for Amazon.com, Box, Cisco Systems, Dropbox, Evernote, Facebook, Google, Microsoft, Mozilla, Nest, Pinterest, Slack, Snapchat, Whatsapp, and Yahoo in support of Apple Inc. at 1-5, 18, In re the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD20, No. CM 16-10 (C.D. Cal. Mar. 3, 2016) (describing the high consumer demand for digital security and consumers' expectations of secure data from major technology companies and their products) [hereinafter Brief of Amici Curiae 1].

9 See, e.g., SYMANTEC CORP., www.symantec.com (last visited Oct. 29, 2016) (Symantec Corp. is a corporation focused on digital security, whereas companies like Apple create encryption and security protocols for their other devices).

10 Cyber Crime, FED. BUREAU OF INVESTIGATION, https://www.fbi.gov/investigate/cyber (last visited Oct. 29, 2016) (“Cyber intrusions are becoming more commonplace, more dangerous, and more sophisticated.”).


*Id.*

*Id.*

*See generally* Brief of Amici Curiae 1, *supra* note 8.


*See id.*

*Id.*

*See id.*

The government has also resorted to requesting that the courts authorize orders compelling suspects themselves to input their own passcodes or otherwise unlock their personal devices. This raises serious constitutional questions, such as a potential violation of the Fifth Amendment right against self-incrimination. This Note focuses on the government compelling companies to provide assistance and does not address individuals being compelled to unlock their devices. For more information on this issue, see RICHARD M. THOMPSON II & CHRIS JAIKARAN, *ENCRYPTION: SELECTED LEGAL ISSUES* 6-16 (Mar. 3, 2016).


The backdoor problem refers to the belief that if a company creates software that can decrypt its encryption software, any party that obtains the software can use it, and so if the software is stolen or otherwise falls into the wrong hands, it can be extremely dangerous. The backdoor problem is covered in more detail in Part I.C.2 Private Actors’ Data Security and “Backdoor” Concerns.


*Id.*


Id.

Id.

Id.

See generally Going Dark, supra note 18.

Motion to Vacate Order Compelling Apple Inc. to Assist Agents in Search and Opposition to Government's Motion to Compel Assistance, In re Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. 5:16-cm-00010-SP, at 2-3 (C.D. Cal. Mar. 3, 2016) [hereinafter Motion to Vacate].


Berger, supra note 36.

Id.

See THOMPSON & JAIKARAN, supra note 22, at 1.


Id.

Brief of Amici Curiae 1, supra note 8.

Epstein Testimony, supra note 25 (citing The Risks of Key Recovery, Key Escrow, and Trusted Third-Party Encryption: Leading Cryptographers and Computer Scientists Report Says Government Encryption Plan is Risky and Impractical) (full citation omitted).

See Brief of Amici Curiae 1, supra note 8, at 18-21.

See id.

Cf. Motion to Vacate, supra note 35, at 23.

Id.

Id.

Id. at 23-24.
52 Id. at 23.


54 The government also has an interest in the international consequences of the United States' ultimate decision on this issue. This Note does not explore these questions or the issues regarding international diplomacy problems regarding device encryption. However, such problems are discussed in Brief Of Amici Curiae Access Now and Wickr Found. In Support Of Apple Inc.'s Motion To Vacate (In the Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD20) (C.D. Cal. Mar. 1, 2016). The government also has an interest in preserving the separation of powers doctrine by ensuring that the correct branch determines the ultimate outcome of this matter and that no branch is overstepping its constitutional authority. This Note will address those concerns and ultimately advocate that the legislative branch is the proper authority to determine this issue in Part III.

55 Going Dark, supra note 18.


57 Id.


59 Id.; see also Tom Risen, Top FBI Attorney Worried About WhatsApp Encryption, U.S. NEWS (Apr. 5, 2016, 5:43 PM), http://www.usnews.com/news/articles/2016-04-05/top-fbi-attorney-james-baker-worried-about-whatsapp-encryption?int=A6909 (FBI general counsel stated that “stronger encryption can also benefit the government.” Risen added that Baker concluded by “noting that data about himself and his family were exposed during a massive Office of Personal Management breach that affected an estimated 21.5 million federal employees or job applicants”).

60 Davis, supra note 58; see also Risen, supra note 59.

61 Davis, supra note 58; see also Risen, supra note 59.

62 Shear & Sanger, supra note 56.

63 See id.

64 See id.

65 Id.

66 Government's Motion to Compel Apple Inc. to Comply with this Court's February 16, 2016 Order Compelling Assistance in Search at 22, In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. CM 16-10 (C.D. Cal. Feb. 19, 2016) [hereinafter Government's Motion to Compel]. https://www.justice.gov/usao-cdca/file/826836/download.

Several months after this matter arose, the FBI claimed that it was eventually able to unlock the iPhone without Apple's assistance and has since dropped its demand. See Devlin Barrett & Daisuke Wakabayashi, FBI Opens San Bernardino Shooter's iPhone; U.S. Drops Demand on Apple, WALL STREET J. , (Mar. 28, 2016, 10:20 PM), http://www.wsj.com/articles/fbi-unlocks-terrorists-iphone-without-apples-help-1459202353. While the parties are no longer arguing about this specific case, the legal issues at the core of this debate are far from resolved. See id. There are still multiple similar cases between Apple and the FBI pending. Furthermore, as reported by the Wall Street Journal, Eric Berg, special counsel for Foley & Lardner communicated that “[T]echnology companies [will continue] to make their devices harder to crack with each iteration. In time, another case will test the issues of privacy and security again ...” Indeed, Berg believes that “‘[i]t's only a matter of time.’” Id.


Andrew Hart, Apple Will No Longer Unlock iOS Devices For Police, HUFFINGTON POST (Sept. 18, 2014), http://www.huffingtonpost.com/2014/09/17/apple-unlock-devices-police_n_5840064.html (explaining the different features included in iOS 8, notably that Apple itself no longer has the ability to decrypt its devices).

Cook, supra note 69.

See Letter from Tim Cook, supra note 42.


See In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by This Court, 149 F. Supp. 3d 341, 350 (E.D.N.Y. Feb. 29, 2016).


Order Compelling Apple Inc. to Assist Agents in Search, In re Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. 5:16-cm-00010-SP, at 21-22 (C.D. Cal. Mar. 3, 2016).

See id. at 25-26; see also Motion to Vacate, supra note 35, at 15-19.

See THOMPSON & JAIKARAN, supra note 22, at 25-26; see also Motion to Vacate, supra note 35, at 15-19.

Motion to Vacate, supra note 35, at 16.

Id. at 17.

See THOMPSON & JAIKARAN, supra note 22, at 26; see also Motion to Vacate, supra note 35, at 15-19.

Motion to Vacate, supra note 35, at 17.

See id. at 15-19.

In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by This Court, 149 F. Supp. 3d 341, 350 (E.D.N.Y. Feb. 29, 2016).

Motion to Vacate, supra note 35, at 17.
Government's Reply in Support of Motion to Compel & Opposition to Apple Inc.'s Motion to Vacate Order at 10, In re Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. 15-0451 (C.D. Cal. Feb. 25, 2016) [hereinafter Government's Reply].

See id. at 11.

Id. at 10-11 (quoting The Company v. United States, 349 F.3d 1132, 1145 n.26 (9th Cir. 2003)).

See id. at 11.

Id. at 11-12.

Id. at 12.

Government's Motion to Compel Apple Inc. to Comply with this Court's February 16, 2016 Order Compelling Assistance in Search at 22, In re Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. CM 16-10 (C.D. Cal. Feb. 19, 2016) [hereinafter Government's Motion to Compel].

See THOMPSON & JAIKARAN, supra note 22, at 2-3.

Government's Motion to Compel, supra note 94, at 23.

See id.

Id.

In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by this Court, 149 F. Supp. 3d 341, 354 (E.D.N.Y. 2016).

THOMPSON & JAIKARAN, supra note 22, at 16.


In re Order Requiring Apple, Inc. to Assist, 149 F. Supp. 3d at 350.


Id. at 178-79.

In re Order Requiring Apple, Inc. to Assist, 149 F. Supp. 3d at 351.


Plum Creek Lumber Co. v. Hutton, 608 F.2d 1283, 1289-90 (9th Cir. 1979).

In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by this Court, 149 F. Supp. 3d 341, 357 (E.D.N.Y. 2016).


Universal City Studios, Inc. v. Corley, 273 F.3d 429, 449 (2d Cir. 2001).

Motion to Vacate, supra note 35, at 32-34.
Government's Reply, supra note 88, at 31-34; see also Brief for Greg Clayborn et al. as Amici Curiae at 18-19, In re Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. 5:16-cm-00010-SP, at 19 (C.D. Cal. Mar. 3, 2016) [hereinafter Brief for Greg Clayborn et al.].

Brief for Greg Clayborn et al., supra note 114, at 18 (quoting Universal City Studios v. Corley, 273 F.3d 429, 454 (2d Cir. 2001)).

Id. (quoting Commodity Futures Trading Comm'n v. Vartuli, 228 F.3d 94, 111 (2d Cir. 2000)).

Id. at 18.

Id. at 8.

Motion to Vacate, supra note 35, at 34.


Id. (internal citations omitted).

Brief for Greg Clayborn et al., supra note 114, at 17 (quoting Sinaloa Lake Owners Ass'n v. City of Simi Valley, 864 F.2d 1475, 1484 (9th Cir. 1989)).

As explained below in this Part, privacy would normally be included as a Fourth Amendment issue in Part II with the other constitutional concerns, but because the AWA requires a search warrant and thereby almost certainly complies with Fourth Amendment, the claim is listed here as more of a practical concern than a legal one. See THOMPSON & JAIKARAN, supra note 22, at 16 n.113.


U.S. CONST. amend. IV.


See Brief for Greg Clayborn et al., supra note 114, at 10-11.

THOMPSON & JAIKARAN, supra note 22, at 16 n.113.

See U.S. CONST. amend. IV (stating “but upon probable cause ...”).


Feldman, supra note 131.


See Brief of Amici Curiae 1, supra note 8, at 10.
In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by this Court, 149 F. Supp. 3d 341, 355-63 (E.D.N.Y. 2016).

See generally id.

Brief for iPhone Security and Applied Cryptography Experts as Amici Curiae Supporting Apple, Inc.’s Motion’s to Vacate Order Compelling Apple, Inc. to Assist Agents in Search, and Opposition to Government’s Motion to Compel Assistance at 22, In re Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203, No. 15-0451 (C.D. Cal. Mar. 22, 2016).

Motion to Vacate, supra note 35, at 18.

Feldman, supra note 131.

Robson, supra note 124 (citing Clark v. Martinez, 543 U.S. 371, 391 (2005) (Thomas, J., dissenting)).

The separation of powers doctrine is a “political doctrine of constitutional law under which the three branches of government (executive, legislative, and judicial) are kept separate to prevent abuse of power. Also known as the system of checks and balances, each branch is given certain powers so as to check and balance the other branches.” Separation of Powers, LEGAL INFO. INST., https://www.law.cornell.edu/wex/separation_of_powers (last visited Oct. 29, 2016).


Feldman, supra note 131.

See In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by This Court, 149 F. Supp. 3d 341, 353 (E.D.N.Y. Feb. 29, 2016) (Judge Orenstein noted there is some gap which Congress has not addressed, and that the question at hand is to determine whether the AWA can fill that gap).

While it is difficult to prove, the government might very well be attempting to “short-circuit public debate on this controversy,” as Apple claims, by using the courts, which should raise suspicion of an attempt to violate the separation of powers doctrine. In re Order Requiring Apple Inc. to Assist in the Execution of a Search Warrant Issued by the Court, 2015 WL 5920207, at *3 n.1.

In re Order Requiring Apple Inc. to Assist in the Execution of a Search Warrant Issued by the Court, 2015 WL 5920207, at *3 n.1.

Chadha, 462 U.S. at 923 (citing Buckley v. Valeo, 424 U.S. 1 (1976)).


Digital Security Commission Act, H.R. 4651, 115th Cong. § 3(b) (2016).


In re Order Requiring Apple, Inc. to Assist in the Execution of a Search Warrant Issued by this Court, 149 F. Supp. 3d 341, 344 (E.D.N.Y. 2016).


Id.


See generally THE SOFTWARE ALLIANCE, supra note 155.

See generally INFO. TECH. INDUS. COUNCIL, supra note 156.

Qualification Standard, supra note 153.

See Morrison v. Olson, 487 U.S. 654, 686 (1988) (citing 28 U.S.C.S. § 596(a)(1)) (Court upheld the Independent Counsel Act, in which a prosecutor appointed to conduct an investigation who alleged wrong-doing by officials could be removed only for good cause).

See Marbury v. Madison, 5 U.S. 137, 153 (1803).

Id. at 177.

Qualification Standard, supra note 153.

See generally National Labor Relations Board: 80 Years of Protecting Employee Rights, NAT'L LABOR RELATIONS BD. (2015), https://www.nlrb.gov/sites/default/files/attachments/basicpage/node-1536/NLRB%2080th%20Anniversary.pdf (illustrating the 80-year history and success of the National Labor Relations Board). Furthermore, according to scholars such as Michael L. Wachter, Professor of Law and Economics at the University of Pennsylvania Law School, “the NLRB has been largely successful and in one key area exceedingly successful.” Michael L. Wachter, The Striking Success of the National Labor Relations Act, in PENN LAW LEGAL SCHOLARSHIP REPOSITORY, RESEARCH HANDBOOKS IN LAW AND ECONOMICS 427 (Cynthia L. Estlund ed., 2012). According to the Board, “The National Labor Relations Board (NLRB) has counted millions of votes, investigated hundreds of thousands of unfair labor practice charges, and issued thousands of decisions.” Graphs & Data, NAT'L LABOR RELATIONS BD., https://www.nlrb.gov/news-outreach/graphs-data (last visited Oct. 29, 2016). Other examples of successful adjudicative agencies with rulemaking authority include the United States Environmental Protection Agency (EPA), the United States Securities and Exchange Commission (SEC), and the Federal Trade Commission (FTC).
When scholars—especially legal academics—talk about intelligence oversight, they typically have in mind a set of processes and institutions designed to deter and detect illegality and abuse. In this Article, I focus on another sense of intelligence oversight and a different institutional actor capable of providing it. The kind of oversight that I describe and endorse is distinguished by its concern with promoting effective intelligence collection while seeking to minimize a wide range of costs, including diplomatic blowback, economic harm to American firms, and intrusiveness that threatens privacy rights. The institution that has begun to furnish this more holistic sort of oversight, and that enjoys conspicuous advantage over preexisting bodies in doing so, is the President, aided by his staff (including those serving on the National Security Council).
Pressured by a constellation of prominent interest group actors, including allied governments and technology firms, the President has begun to weigh in on surveillance policy and to shape the metes and bounds of intelligence collection in a systematic fashion. This development—which I call presidential intelligence—bears a family resemblance to presidential administration, the turn to centralized, political control that has dominated the scholarship and practice of administrative law for over a generation. In this Article, I offer a descriptive account of the rise of presidential intelligence, a qualified normative defense of its value (as an addition to, rather than a replacement of, existing oversight bodies), and a set of prescriptions for how to design institutions in order to realize its full potential.

**INTRODUCTION**

For a generation we have “live[d] ... in an era of presidential administration.”¹ Whether exercising power directly or through White House units like the Office of Information and Regulatory Affairs (OIRA), Presidents of both parties, employing a variety of mechanisms and summoning a range of justifications,³ have sought to leave an imprint on the regulatory state.⁴ Presidential administration serves not only as a font of centralized power and control, but also as a source of democratic accountability for an administrative state perennially anxious about its legitimacy. The tectonic shift toward presidential control of agencies has reverberated throughout the federal bureaucracy, including a large swath of the national security state—ewith the striking exception of the so-called “intelligence community.”⁶

A major reason for intelligence's exceptionality is historical. In the aftermath of Watergate and the intelligence scandals exposed by the Church and Pike Committees, the reigning assumption was that, of the three branches of government that might exercise meaningful oversight of the intelligence apparatus, the possibility of heightened presidential authority ought to be taken off the table. That is because the architects of the new oversight took presidential control as a given and saw in the White House-intelligence complex the capacity for tyranny and abuse. To resist executive dominance, they chose to empower the other branches of government and to interpose a range of traditional, as well as internal, separation of powers checks. These checks include the Foreign Intelligence Surveillance Court (FISC), specialized congressional oversight committees, inspectors general within the various agencies, and the Intelligence Oversight Board (currently organized as a component of the President's Intelligence Advisory Board (PIAB)). Leading national security law scholarship has also characteristically regarded presidential abuse as one of its points of departure. From Professor Harold Koh's pioneering work 25 years ago to Professor Jack Goldsmith's recent emphasis on the role of “soft law” norms and civil society institutions in constraining the national security executive, scholars have tended to assume that the White House is and ought to be an object, not a source, of intelligence oversight. To practitioners and scholars in this area, the idea of entrusting the President to oversee intelligence is deeply counterintuitive.

But the arrangement seems strange only because of a basic misconception that the intelligence community marches in lockstep with the White House. It does not.¹⁶ In fact, a decentralized intelligence community that has proved adept at empire building and has been largely unconstrained by the political executive has revealed itself to be profoundly vulnerable to questionable intelligence-gathering practices. Indeed, the intelligence community has carried out a range of activities that, while conferring uncertain benefits, have led to significant diplomatic blowback, jeopardized the bottom lines of American industry, and pushed the envelope (at the very least) on questions of privacy and civil liberties.
What I refer to as presidential intelligence—the White House's sustained, routinized, and process-driven governance of American spying—takes as its starting point these misalignments between the political executive and the intelligence community. Presidential intelligence seeks to address such misalignments by harnessing the White House's unique capacity to shape the metes and bounds of intelligence collection in a systematic, ongoing way. Presidential intelligence is not merely a good idea; it is an emerging reality on the ground. In this Article, I set out to describe and defend its recent arrival on the scene, and to encourage its future growth through sound institutional design.

I make four main scholarly contributions. First, I show that, as a descriptive matter, the norms of presidential control that have characterized the majority of the regulatory state for decades have recently begun to take hold in the domain of intelligence collection. Transposing the concepts and architectures of presidential administration to national security, and in particular to the world of intelligence, may seem odd, but is essentially plausible. In fact, although they clearly rest on different constitutional foundations, there is a lot to recommend the analogy between the intelligence apparatus and the administrative state, beginning with a shared pedigree: Both are mid-twentieth-century transplants to Washington and are uneasy fits with the preexisting traditions and institutional life of American constitutionalism. Both initially grounded their legitimacy in claims of politically neutral technocracy before later broadening their foundations to rely heavily on legal institutions and processes. But over the last generation, their trajectories have diverged considerably. The intelligence apparatus has in large measure retained its strong ideals of bureaucratic independence from politics, even as the balance of the regulatory state has been transformed by ever-increasing presidentialization. The reabsorption of the intelligence state into the mainstream of administrative law and regulation through its own belated presidentialization is powerful proof not of the exceptionality of national security but rather of its banality.

My second main contribution, also descriptive, is to offer an account of how and why the current moment has proved especially propitious for the ascendancy of presidential intelligence. In particular, I call attention to the role that technology and telecommunications firms on the one hand, and allied governments on the other—all themselves intelligence collectors and connoisseurs—have played in catalyzing and shaping the emerging dynamics in this area. Such actors have made common cause with more traditional civil liberties groups in pushing back against a range of intelligence-collection practices. The emergence of separation of powers-type checks furnished by businesses and foreign governments is itself underwritten by heightened levels of transparency and what one scholar has dubbed “the declining half-life of secrets,” a development that has altered the incentives that used to operate in this area.

My third contribution is to offer a qualified normative defense of the turn to the institutional presidency, especially key White House elements such as the National Security Council (NSC) (but also less conventional actors in national security, like the National Economic Council (NEC) and the Office of Science and Technology Policy (OSTP)), as a source of political direction and accountability for the post-9/11 intelligence bureaucracy. The normative claim itself has three components. First, I argue that presidential intelligence has the capacity to bolster the policy and economic grounds for intelligence decisions. This “strategic turn” entails a reconceptualization of the purpose of intelligence oversight. Ever since the 1970s-era reforms, the assumption—coded into the DNA of oversight bodies—has been that the main task of intelligence oversight is to detect and deter illegality and abuse. Presidential intelligence takes that foundation as a given and seeks to add to it mechanisms designed to promote strategically sound intelligence collection. Second, presidential intelligence has the capacity to enhance the ways in which the intelligence apparatus is made democratically accountable, through a range of institutions and methods (each of them inevitably limited). Third, under many (though not all) specifications, presidential intelligence may enhance certain rights protections—in some cases more effectively than other oversight tools.
My fourth scholarly contribution is to suggest sound institutional design to help realize the potential of presidential intelligence. I call for a mixture of centralized review based in the White House and greater numbers of political appointments (with and without Senate confirmation) in the intelligence agencies.

Presidential intelligence is intended as a complement to existing oversight mechanisms, not a substitute for them. It is certainly not a panacea, any more than presidential administration has proved to be one. I offer no predictions as to where intelligence policy will come to rest in the United States in the coming years, or as to how presidential intelligence will fare in practice if (as I expect) it develops into a defining feature of intelligence governance. But designed smartly, presidential intelligence represents a meaningful opportunity to enhance the effectiveness, accountability, and attentiveness to civil liberties of a crucially important and inevitably delicate instrument of American power.

The structure of the Article is as follows. In Part I, I aim to contextualize the recent ascendency of presidential intelligence within the broad range of presidential involvement in intelligence matters. While my approach is necessarily schematic, I aim to show that White House involvement has been systematic and sustained in certain areas but not others. The historical baseline includes significant presidential involvement in some aspects of the intelligence process (consuming the products of intelligence analysts and governing covert action), moderate involvement in others (strategic agenda setting and budgeting), and only modest involvement (until very recently, as I show) in the oversight of what amounts to the core business of the intelligence community: intelligence gathering. Concerning this latter function, conditions unique to the intelligence enterprise--and the distinctive political incentives that sustained them--have, until recently, cut against presidential involvement. Intelligence collection's defiance of the centripetal forces that dominate American public life across so many domains largely persisted even after 9/11. The so-called President's Surveillance Program (PSP) was assuredly an assertion of White House control over certain collection programs, but, insofar as it amounted to an end run around institutions, process, and law, it lacked the core attributes of presidential intelligence. The Office of the Director of National Intelligence (ODNI), a manifestation of centralization within the intelligence bureaucracy, meanwhile, has suffered from an opposite flaw: lacking in White House backing, its considerable institutional potential has not been realized. All along, while White House involvement has been muted, oversight of collection has been dominated by the “legalist” architecture and ethos that grew up (outside and inside the executive branch) over the last generation, which has tended to be both insufficient to the realization of rights-protective intelligence and ill-equipped as a mechanism for generating sound intelligence policy.

What 9/11 did not achieve in terms of fundamentally reshaping the presidential outlook on intelligence gathering, the Snowden leaks and their aftermath seem poised to accomplish. In Part II, I focus on the changes that are afoot, turning to a detailed description of why and how a President who committed, after his election, to “[g]etting [p]olitics out of [i]ntelligence” has, in fact, taken powerful steps in the other direction. The President has been thrust into the role of overseer concerning a range of intelligence activities carried out by multiple agencies. The emergence of presidential intelligence has been catalyzed--and entrenched--by a new political economy in which telecommunications and technology firms that were explicitly revealed by Snowden to be National Security Agency (NSA) partners, as well as overseas allies that were shown to be NSA targets, have joined privacy advocates in putting pressure on the White House to cut back on certain intelligence-gathering practices. As the President acknowledged in a major policy address that debuted the makings of a new White House-based oversight regime, “I've listened to foreign partners, privacy advocates, and industry leaders.” While the evidence of the emergence of presidential intelligence is not yet overwhelming, I document the significant steps taken in Presidential Policy Directive 28 (PPD-28), the January 2014 order issued by the White House (accompanying the aforementioned address), as well as certain even more recent
developments that point in the same direction. It is perhaps telling that Professor Cass Sunstein, one of the leading
scholars of the regulatory state and a former Administrator of OIRA (an office which he has dubbed “the
cockpit of the regulatory state”), served on a key presidential committee that issued recommendations about how
to move forward after the Snowden leaks. Thus far, the changes have largely conformed to political scientist Professor
Terry Moe's category of centralization (presidential control based at the White House) rather than his concept of politicization (presidential control made effective through appointments in the agencies themselves).

Drawing on the literature of presidential administration—in particular, the scholarship on the White House's direct
involvement in regulation, as well as the academic commentary on the role of OIRA--Part III offers a (necessarily
preliminary) assessment of the prospects, pro and con, of presidential intelligence. On the positive side of the ledger, the
academic literature on OIRA points to two significant upsides of presidential intelligence: an intelligence apparatus
in which the competing interests of multiple agencies are harmonized, and one in which some kind of cost-benefit analysis
(nontechnical and nonmonetized) of proposed intelligence gathering is carried out. A third key aspect of OIRA practice—quantifying the costs and benefits of a proposed major rule in dollar figures—is less readily translatable to the intelligence
arena. Meanwhile, although democratic accountability of the intelligence state will always run up against the limitations
imposed by secrecy, more visibility and revelation of the sort that presidential intelligence entails represents a welcome
improvement on this front. Here, too, the presidential administration literature establishes a conceptual framework and
a standard that presidential intelligence ought to be able to strive for, if not fully achieve. Finally, presidential
intelligence might well promote intelligence that is more attentive to basic rights. Here I return to the role of technology
firms as a de facto pressure group for privacy protections. While these firms are principally motivated by the market,
not morals, their commitment to privacy is now sufficiently baked into the global business strategies they pursue (and
even the devices they bring to the market) that it is likely to prove durable. And even if their antisurveillance sensibilities
eventually give out under a new set of economic pressures, presidential intelligence (and the deliberative process on which
it depends) will have already taken on an institutional life of its own.

Presidential intelligence also carries certain risks. I shed light on three potential downsides in particular. The first is
that presidential control will “politicize” intelligence in the sense of distorting analytic findings in order to placate
policymakers with parochial agendas. This potential weakness—discussion of which is a staple of intelligence studies—
can profitably be thought of as a species of concern that overhangs all administrative law: how to strike the right
balance between technocratic detachment and expertise on the one hand, and political control on the other. While
concerns about distorting expert judgment are well taken, they should not doom the enterprise of presidential intelligence
any more than they undermine the rationale for presidential administration. Second, I contend with the prospect that
presidential intelligence might give impetus to unhealthy institutional dynamics between the White House and Capitol
Hill, exacerbated by (and potentially also fueling) the contemporary phenomenon of hyperpartisanship. Here, too, the
presidential administration literature offers useful context and, if not cause for optimism, then at least some reason to
think that presidential intelligence will not fare any worse on this dimension than presidential administration. Third, I
take up a potential vulnerability unique to the intelligence environment—namely that under certain specifications, fusing
presidential power with intelligence capabilities might tend to recreate the conditions for abusive practices of the sort that
prompted the significant intelligence reforms of the 1970s or that doomed the PSP a decade ago. This last worry has no obvious equivalent in the presidential administration repertoire. It can be ameliorated, if not ultimately resolved,
through a renewed commitment to external checks—prominently including judicial review of intelligence matters.

The prospects of presidential intelligence depend not only on macro-level trends, but also on carefully crafted
institutional design. In Part IV, I consider two specific approaches that have the potential to stimulate growth in the
right direction. First, I take up a presidential “finding” requirement for certain key collection programs or practices, on
par with the requirement that the President sign off before covert action is undertaken. Second, I offer support for a more thoroughly politicized (in Moe's sense) intelligence bureaucracy, with greater numbers of presidentially nominated (and potentially also Senate-confirmed) senior officials. Among other things, ramping up political leadership within the intelligence agencies can help counteract the tendency--much commented on of late--of the White House to cross the line from centralized control to micromanagement of the bureaucracy, including (or especially) in matters of national security. Some of these changes are achievable without new legislation, most likely through amending Executive Order 12,333, which governs the intelligence community.

I conclude by contemplating what an intelligence community that has been absorbed into the heartland of the regulatory state-- rather than treated as legally and intellectually quarantined from the balance of public law and policy--will look like. Having demonstrated throughout the Article how the scholarly literature on, and the practical experience of, the administrative state can and should inform our understanding of national security law and policy, I suggest some ways in which the gains of trade might flow both ways, pointing out how concepts well known within the precincts of intelligence practice and scholarship might illuminate problems in “ordinary” administrative law.

*646 I. THE PRESIDENT AND THE INTELLIGENCE COMMUNITY: A BASELINE

As I discuss below, the relationship between the President and intelligence collection has largely defied the logic of presidential control. In order to contextualize the recent changes in this area, it is important to establish a rough baseline of where presidential intelligence stood prior to the developments that I analyze, and how the President has historically related (or failed to relate) to the intelligence state across a wide range of intelligence practices. What emerges is that the scope of presidential involvement in different aspects of intelligence has varied extensively. Certain areas-- consumption of intelligence analysis and management of covert action--are highly presidentialized (albeit for quite different reasons). Bureaucratic management of the intelligence agencies--including agenda setting and fiscal oversight--has been somewhat presidentialized. Meanwhile, until the recent developments I document later in this Article, the White House has been largely absent from the systematic oversight of how intelligence is collected--the bread and butter of what spy agencies do.

A. Analysis and Covert Action: Highly Presidentialized

The President has always been the consumer-in-chief of intelligence, and the intelligence community has always stood prepared to advise him (and his senior staff) on issues of concern. Each day, the intelligence community prepares an intelligence digest for the President that is then briefed to him, in person, by the Director of National Intelligence (DNI) (or his designate). The President is able to probe, demand follow-up on issues, and shape the way in which intelligence is presented. The President's incentive to pay attention to his intelligence briefing is essentially strategic: he ignores it and its implications for American security and foreign policy (and, ultimately, his own political standing) at his own peril. This is not to suggest that the President is able to dictate substantive conclusions. Within the analytic arms of the intelligence bureaucracy, including in offices like the National Intelligence Council, there is a powerful cultural sensibility that militates in favor of neutrality and a “speak-truth-to-power” ethic. But the President has extensive contact with intelligence analysts and analysis through the daily briefing he and other senior policymakers receive, and he is able to exert influence over the substantive agenda of the intelligence community (including, at least indirectly, over collection) by pointing it toward--or away from-- specific areas.
Another critical node of intense presidential involvement in intelligence involves covert action.\(^5\) In the first decades of post-World War II intelligence, Presidents were leery of proximity between the White House and CIA operations,\(^5\) and intelligence officials were, for their part, happy to respect presidential preferences for plausible deniability. This bargain came undone in response to the long shadow of Watergate, as well as the significant findings of malfeasance painstakingly documented by the Church Committee.\(^6\) Intelligence reformers in 1974 imposed the requirement that “no appropriated funds could be expended by the CIA for covert actions unless and until the President found that each such operation was important to national security, and provided the appropriate committees of Congress with a description and scope of each operation in a timely fashion.”\(^7\) In other words, the oversight solution for the problem of covert action has been to restrict the availability to the White House of plausible deniability and to compel the President to determine whether a course of action is, on balance, worth the risk. As discussed below, academics debate how robust the presidential finding process is in application.\(^8\) Meanwhile, *648* public debate about recent covert-action programs paints a complex picture, with the Senate Select Committee on Intelligence report suggesting that the CIA misled the White House on aspects of its detention and interrogation program,\(^9\) even as President Obama has been characterized as being intimately involved in the drone program.\(^10\) In general terms, though, the statutory demand for White House involvement has been a vector for reshaping the President's incentives on covert action.

### B. Organization and Budget: Somewhat Presidentialized

The President enjoys considerable authority to shape the intelligence community from the standpoint of its structure, budget, and organizational priorities. As set out in Executive Order 12,333, the charter order that has governed the intelligence state for over thirty years,\(^11\) the President is empowered to specify the roles and responsibilities of various components of the intelligence community.\(^12\) This executive order looms especially large because statutes governing the allocation of power to intelligence agencies are notoriously vague,\(^13\) or *649* in some cases nonexistent.\(^14\) Within the White House, presidential management of the intelligence bureaucracy is pegged to the key--if often unsung--leadership role assigned to the Assistant to the President for National Security (known popularly as the National Security Advisor) by Executive Order 12,333.\(^15\) The National Security Advisor maintains a small, dedicated intelligence staff on the NSC, typically headed by a senior intelligence official.\(^16\)

The White House also has access to the President's Intelligence Advisory Board (a subset of which also functions as the Intelligence Oversight Board), originally conceived by President Eisenhower as a means of furnishing “advice to the President concerning the quality and adequacy of intelligence collection, of analysis and estimates, of counterintelligence, and of other intelligence activities,” though scholars debate its overall utility.\(^17\) In addition, the President has the power to convene ad hoc commissions and task forces. In the aftermath of the Snowden leaks, President Obama actually convened three--one, known as the President's Review Group, was made up of five members with legal and intelligence backgrounds.\(^18\) Another was *650* staffed within the NSC itself.\(^19\) And a third was focused on the relationship between “big data” and privacy.\(^20\)

The White House's capacity to shape the agenda of the intelligence bureaucracy in terms of what “requirements” they collect against is also considerable. For example, President Clinton issued Presidential Decision Directive 35 to establish intelligence priorities in a post-Cold War landscape.\(^21\) And in National Security Presidential Directive 26, the George W. Bush Administration provided guidance to the intelligence community by creating “a dynamic process for articulating and reviewing intelligence priorities.”\(^22\) The National Intelligence Priorities Framework (NIPF) was
established to implement this directive and translate White House priorities into concrete deliverables. The NIPF process offers intelligence “customers” across the national security state an opportunity to rank their various priorities for the intelligence agencies, and then affords the National Security Advisor (working through the interagency process at the NSC) the ability to come up with an overall recommendation to the President. Intelligence professionals complain that the process suffers from a one-way ratchet problem--intelligence requirements never come off the list or get explicitly downgraded in priority. They are merely eclipsed by new priorities. But whatever its limitations, the NIPF provides the White House a significant say in shaping strategic priorities for intelligence.

While the intelligence budgeting process remains opaque and involves bureaucratic sleights-of-hand like “reprogramming,” it is clear that the White House, with the assistance of a small, dedicated intelligence staff at OMB, plays a key role here. The President's power stems, in part, from the way that the intelligence budget is divided into two components: a National Intelligence Program (NIP), which is managed by the Director of National Intelligence, and a Military Intelligence Program (MIP), which is overseen by the Secretary of Defense. For Fiscal Year 2015, the aggregate amount of appropriations requested for the NIP was approximately $50.4 billion while the request for the MIP budget was approximately $16.6 billion. Concerning both, notional budgets are prepared by the various component agencies, and are then rationalized and coordinated by the DNI and the Secretary of Defense. But it is only at the White House that the two budgets are harmonized before being sent to the relevant committees on Capitol Hill.

C. Intelligence Collection: Weakly Presidentialized

All of the aforementioned points of contact between the White House and the intelligence agencies are, of course, hugely important. But at the heart of my argument is the fact that when it comes to the sustained oversight of how intelligence is collected--what has rightly been called “the bedrock of intelligence” -- the President's role has been limited. Certain highly sensitive collection programs do garner White House attention. But it remains the case that the core “business” of spy agencies (running the gamut from the FBI to the CIA to the NSA, and so on) is largely ungoverned by the White House. Intelligence scholars have long called for tighter political control of intelligence collection, such as when Professor Harry Howe Ransom recommended that “[n]o foreign secret action should be undertaken until after the most careful weighing of risks against possible gains, and particularly a careful and realistic analysis of the prospects for secrecy and the consequences of public exposure.” But unlike the case of covert-action regulation, there has been (at least until Snowden) no watershed culminating in a formal demand that Presidents pay systematic attention to intelligence gathering. To understand why, it is useful to step back and consider the sorts of centripetal forces that operate across the broad sweep of American public life, and to analyze why they have tended not to apply in this area.

Although the “institutional presidency” is an artifact of the New Deal, and the Executive Office of the President came about in 1939, positive political science theories until recently tended to focus on the workings of Congress at the expense of the presidency. But Moe and Scott Wilson accurately perceived that the conditions of modern political life favored an ever-growing role for Presidents because presidential incentives to intervene in any given issue were driven by the politics of accountability. As Moe had previously argued, “the expectations surrounding presidential performance far outstrip the institutional capacity of presidents to perform. This gives presidents a strong incentive to enhance their capacity by initiating reforms and making adjustments in the administrative apparatus surrounding them ....” When it comes to the assertion of control over the bureaucracy, two strategies are available to Presidents. First, through a process of “centralization,” “presidents can move toward coherent central control by setting up their own policymaking structures inside the White House, incorporating people of their own choosing from the departments, the agencies,
and the Executive Office, and pulling salient issues of public policy into the presidency for debate and resolution.”

The creation of regulatory review within OMB under President Reagan was an important manifestation of this sort of centralizing impulse. So was President Clinton's more personal--and more selective--involvement in certain signature policy arenas documented by then-Professor Elena Kagan in her landmark article that conveyed Moe's insights into the heartland of administrative law. The second strategy for the President to assert control over the bureaucracy is “politicization,” which entails the assertion of presidential control through the appointment of officials in the agencies in order “to ensure that important bureaucratic decisions are made, or at least overseen and monitored, by presidential agents.” As then-Professor David Barron has put it, “for the last three decades, Presidents have been doing much more than looking for ways to wrest discretionary decision-making power from agencies. Over that same period of time, Presidents also have been making novel and aggressive use of their powers of appointment to remake agencies in their own image.”

Returning to the intelligence orbit, and the absence of sustained, routinized presidential involvement in intelligence collection, it is necessary to consider the preexisting constellation of incentives and institutional dynamics. First, up until recently, there was no obvious incentive for the President to superintend intelligence collection methods akin to the President's incentives to pay close attention in his daily consumption of intelligence analysis. Quite the opposite, the President has, over the years, come to embrace his marginal status as overseer of intelligence-gathering methods. Intelligence oversight, historically a secret means for governing secret programs and processes, offered few reliable political benefits. As intelligence scholar Professor Amy Zegart has observed:

> Intelligence is in many respects the worst of all oversight worlds: It concerns complicated policy issues that require considerable attention to master, deals with highly charged and controversial policies that are fraught with political risk, requires toiling away in secret without the promise of public prestige, and provides almost no benefit where it counts the most, at the polls.

Concerning secret intelligence collection, the public would not blame the President because there would be no visible problem for which blame needed to be assigned. At the same time, for a President to become enmeshed with the management of intelligence collection carried potential downsides, including the always salient opportunity costs measured in efforts not expended on issues offering political rewards, as well as potential proximity to a set of activities that, by design, push the outer limits of legality. If the President could not benefit politically from investment in oversight of collection, and if he stood to lose a great deal, it would obviously be advisable to cede ground on oversight issues to the coordinate branches of government. Certainly the intelligence agencies themselves (which acted as an unopposed interest group of sorts in this domain) never sought out greater supervision from their reluctant political principal. Reinforcing the incentives of presidential noninvolvement was a lack of institutional mechanisms and legal prompts for focusing White House attention. As noted above, two of the most important mechanisms for engaging the President on intelligence matters—the President's daily brief and the covert-action finding—did not speak directly to the oversight of collection.

The post-9/11 counterterrorism imperative, for all the profound changes in national security policy, law, and institutional design that it ushered in, did not fundamentally recast the President's relationship to intelligence collection. The well-known story of the President's Surveillance Program (PSP), codenamed “STELLAR WIND,” confirms this point. Although the President and some of his senior advisors were directly involved in authorizing the PSP, that involvement did not amount to a manifestation of presidential intelligence in the sense that this Article employs the
concept. First, and most fundamentally, presidential intelligence as I conceive it depends on the existence of sustained, routinized governance by the White House and its components, especially the NSC. In this respect, the PSP was the antithesis of presidential intelligence. It was the product of an intensely secretive effort that was unable to withstand legal scrutiny from within the executive branch and effectively collapsed under its own weight when aspects of the program came to light. Furthermore, the secrecy with which the PSP was carried out, including within the government itself, is not reflective of the baseline assumptions that have lately catalyzed and shaped presidential intelligence. Nor did the PSP feature the sorts of internal processes that are central to the enterprise of presidential intelligence. It was certainly not marked by thorough assessments of risk that took into account far-reaching factors such as economic and diplomatic fallout. It was only when the Bush Administration made its case for aspects of the program to the entire Congress in 2007 and 2008 that the sorts of tradeoffs inherent in electronic surveillance at scale, such as its potential impacts on telecommunications firms (which sought and obtained immunity from Congress), were considered.

Nor did the creation of the ODNI, the result of a recommendation of the 9/11 Commission, signify presidential control. A superagency meant to sit atop the entirety of the intelligence bureaucracy, the ODNI only fortified the independence of the intelligence community by interposing another layer of bureaucracy between the President and the spy agencies. This was partly by design; President Bush fatefully rejected a proposal that would have lodged the DNI within the White House. The statute also reflected a bitter compromise between intelligence reformers and backers of the Pentagon, such that the DNI's authority was sharply curtailed from the start. It hasn't helped that in its first decade there has been high turnover in the top job, or that DNIs have been repeatedly outmaneuvered by bureaucratic rivals--most notably when then-CIA Director Leon Panetta managed to scuttle DNI Admiral Dennis Blair's effort to dislodge the CIA from its historic role as the lead intelligence presence in American embassies throughout the world. The fight was ultimately adjudicated by the Vice President in favor of the CIA. Although the DNI is himself picked by the President and confirmed by the Senate, the organization as a whole seems to have foundered without significant White House ties and prestige. Intelligence scholar Professor Loch Johnson is too harsh in deeming the DNI a “cardboard cutout,” but his criticism is not baseless.

To say that centralized political control of intelligence collection has generally been weak is not to suggest that intelligence gathering has been altogether impervious to oversight. Legalist institutions and processes have played an important role, helping to determine “whether [the] Government's intelligence activities [are] governed and controlled consistently with the fundamental principles of American constitutional government” and to interpose “effective measures to prevent intelligence excesses.” Institutionally, legalist oversight is typically associated with separation of powers checks like the FISC and congressional intelligence committees. But in practice, much of the checking for compliance with law takes place within the executive branch through a wide range of institutional actors. Offices of general counsel within the various intelligence agencies have swollen in size and institutional significance over the last generation. The NSA now features an office of compliance, and a number of intelligence agencies have dedicated civil liberties officers. Inspectors general wield substantial power across the intelligence state. The Department of Justice also plays a pivotal role in providing legalist oversight, through everything from Office of Legal Counsel (OLC) memos, to the formulation of guidelines for FBI surveillance, to the role of the Attorney General in authorizing certain types of electronic surveillance as required by the Foreign Intelligence Surveillance Act (FISA) and in approving each intelligence agency's internal regulations for collection pursuant to Executive Order 12,333. Finally, quasi-independent entities have figured in legalist oversight. The (largely moribund) Intelligence Oversight Board nominally superintends compliance with a wide range of legal constraints, while the bipartisan Privacy and Civil Liberties
Oversight Board (PCLOB)\textsuperscript{123} has recently assumed an important role \textsuperscript{658} in legalist oversight of counterterrorism measures. \textsuperscript{124}

And yet, as I elaborate below, legalist oversight of intelligence collection is no substitute for presidential control. This is partly because the ambition of legalist oversight is considerably narrower; it does not aspire (and in any event is not equipped) to engage with the strategic wisdom of intelligence gathering. But even when it comes to attending to privacy and civil liberties, existing legalist controls are wanting, as Professor Margo Schlanger has recently argued. \textsuperscript{125} This is due to questionable legal interpretations, \textsuperscript{126} flawed compliance practices, and imperfectly designed oversight bodies like the FISC, which has struggled with a lack of in-house technical expertise, the absence of an adversarial structure, \textsuperscript{127} and insufficient “capacity to investigate issues of noncompliance.” \textsuperscript{128} Furthermore, when it comes to the large swaths of intelligence collection governed by Executive Order 12,333--which dwarf the amount of collection under FISA--legal controls are generally less robust, predicated as they are on internal agency regulations promulgated under the order. \textsuperscript{129} For this reason, to a growing number \textsuperscript{659} of scholars and privacy activists, collection under 12,333 represents the next frontier of intelligence controversy. \textsuperscript{130}

II. THE EMERGENCE OF PRESIDENTIAL INTELLIGENCE

While the emergence of presidential intelligence does not lend itself to precise periodization, it is my contention that the revelation of surveillance practices by Edward Snowden can be thought of as the fulcrum of the transition. Goldsmith captures the moment and its upshot with precision:

\begin{quote}
Pre-Snowden, the US government faced few constraints in its collection and analysis other than what the law imposed, what its technology could achieve, and what its large budget permitted. Within these constraints, it could focus solely on the national security benefit side of communications surveillance, for there were few costs, and practically no political costs, to it. In the post-Snowden world, NSA collection programs are very costly along many dimensions, and the US government faces many tradeoffs and conflicting interests. \textsuperscript{131}
\end{quote}

Under conditions of unprecedented visibility, political blame was assigned to the White House for perceived intelligence excesses and the President was compelled to assume greater control of the issues. The particular causal mechanisms that prompted presidential intelligence to take hold are inevitably numerous and overlapping.

The Snowden leaks themselves (which came on the heels of the Julian Assange revelations and have already been followed by others), as well as the White House attention these leaks generated, can be seen as a predictable response to the exponential growth that the intelligence bureaucracy has undergone since 9/11. Massive growth has increased the risk surface for leaks, both in the sense that it has become that much harder to ensure the impenetrability of a greatly enlarged and complex workforce\textsuperscript{132} (Snowden was working as a contractor) and \textsuperscript{660} because the expanded ambitions of post-9/11 intelligence have created more potential points of friction that could, in turn, galvanize insiders to expose what they perceive as official excess. Indeed, thinking even more macroscopically about the nature of technology itself, it has both significantly expanded the capacities of intelligence agencies to collect information and massively increased vulnerability, for example by empowering individuals within the intelligence apparatus to expose official practices on a heretofore unimaginable scale. Professor Peter Swire has convincingly argued that secrets have a “declining half-life” and that intelligence agencies fail to internalize this reality at their own peril. \textsuperscript{133} Swire's view appears to have shaped one of
the recommendations of the President's Review Group on which he served: “[W]e should not engage in any secret, covert, or clandestine activity if we could not persuade the American people of the necessity and wisdom of such activities were they to learn of them as the result of a leak or other disclosure.” While all of these structural accounts undoubtedly contain explanatory power, my account emphasizes the ways in which recent revelations (and the ongoing, realistic prospect of more) have unleashed new patterns of interest group involvement in this area. Regardless of the complex reasons for its emergence, presidential intelligence is not likely to fade from the scene when the Snowden moment has come and gone.

A. The New Political Economy of Intelligence

Since 9/11, there has certainly been interest group contestation in national security, with civil liberties groups tending to oppose a range of government policies that they view as privileging security over core constitutional rights of expression, liberty, due process, and privacy. But what is distinctive about the post-Snowden developments is that other, arguably more powerful, groups have united with privacy activists to challenge official surveillance policy. This is not to devalue the work of privacy groups, which themselves have proved technologically capable and politically nimble. Still, the fact that civil liberties advocates have resisted intelligence practices of late is hardly news; they have done so with mixed success for well over a generation. The business firms and foreign governments who have now added their voices to the discussion bring considerable economic and diplomatic clout to the table, as well as sophistication about intelligence, being connoisseurs and practitioners themselves. It had previously been written of intelligence that “[f]ew interest groups exist in this policy domain.” But the Snowden revelations helped to usher in a change on this front. Under pressure from this new constellation of actors, the White House has been forced to recalibrate its own outmoded assessment of the relative costs and benefits of disengagement from intelligence governance.

1. Technology Firms and Economic Misalignment.--A critically important--and thus far, largely unheralded (at least by scholars)--feature of the new intelligence oversight ecosystem is the role of American technology and telecommunications firms. These firms combine deep understanding of the nature of signals intelligence--much of the work that the intelligence community performs is done collaboratively with private actors--with sensitivity to the global marketplace and worries about reputational and economic harms that could result from being identified with the putative misdeeds of the NSA. As Julian Sanchez has put it, “perhaps the most significant change wrought by the Snowden disclosures to date has not been the policy proposals it has inspired--which, however vital, tend to focus on rules rather than architectures--but in the way it has transformed the incentives of the technology companies that maintain those architectures.”

The technology firms have certainly been outspoken on these matters. For example, following recent reports that the NSA was able to access information in the custody of U.S. technology companies outside the United States through means not supervised by the FISC, Google's Chief Legal Officer, David Drummond, said that the company was “outraged” that the government would have intercepted data from Google's private networks, which he said “underscores the need for urgent reform.” The message has been received by leaders of the intelligence community. NSA Director Admiral Michael Rogers recently acknowledged that, in the words of his interviewer David Sanger, “the quiet working relationships between the security agency and the nation's telecommunications and high technology firms have been sharply changed by the Snowden disclosures--and might never return to what they once were in an era when the relationships were enveloped in secrecy.” And the firms have certainly registered their displeasure with the surveillance status quo by lobbying Congress to amend public laws. But they have also lodged their protests...
directly with the White House. The President has heard repeatedly from angry technology and telecommunications CEOs who have pressed him for profound changes to surveillance practices. 150

There are some suggestions that the economic fallout from the reputational harm of being closely identified with the intelligence apparatus has already been significant. A recent report cites estimates of the economic losses to American cloud-computing firms owing to global concerns about NSA spying ranging from $22 billion to $180 billion. 151 The Snowden revelations have recently caused the German government to transfer an important contract from Verizon to Deutsche Telekom. 152 The changed economic dynamics have also supplied the backdrop for legal showdowns, 153 such as the unfolding *664 contest in the Second Circuit about a warrant issued to Microsoft to turn over information that the company is storing in Ireland to American prosecutors. 154 The companies are also beginning to engage in forms of commercial “self-help,” employing default encryption technologies on mobile devices and explicitly marketing them as being impervious to government snooping. 155 The message is clear: the global marketplace demands consumer technology (or cloud-based services) that defeats surveillance, and if the Apples of the world are not poised to provide it, some other corporation will. In yet another unmistakable nod to the imperatives of global competitiveness, Google's top lawyer has recently argued for the extension of American privacy protections to EU citizens. 156 In sum, a major American industry has now taken a stance against “overregulation” by the intelligence state--possibly the first time in the annals of post-World War II American national security that a set of powerful economic actors has been so misaligned with the security apparatus and so vocal about it.

Of late, some national security officials have begun to push back. FBI Director James Comey has publicly argued that the pendulum has now swung too far in the direction of privacy, 157 specifically decrying the recent push toward encryption and warning that “Apple and Google have the power to upend the rule of law.” 158 If anything, the backlash in the United Kingdom has been even more pronounced. *665 The recently tapped head of Government Communications Headquarters (GCHQ), the United Kingdom's electronic surveillance arm, publicly castigated American social media platforms for serving as “the command-and-control networks of choice for terrorists and criminals, who find their services as transformational as the rest of us.” 159 But it remains unclear what effect, if any, these interventions from security officials may have. The firms have clearly indicated that they will maintain pressure on officials-- up to and including the President--to rein in what they decry (at least for public consumption) as the excesses of the intelligence state. Comey's recent announcement that the Obama Administration would not seek to legislate “backdoors” to defeat encryption 160 suggests that the technology firms have, at least for now, gained the upper hand.

2. Allies and Strategic Misalignment.--The President has also had to absorb pushback from allies. 161 U.S. envoys were summoned by the French, 162 German, 163 and Brazilian 164 authorities, among others, to explain U.S. surveillance practices, including surveillance of heads of state. No case was more inflammatory than the revelation that the United States had carried out surveillance of German Chancellor *666 Angela Merkel's cellphone. 165 As Henry Farrell and Abraham Newman have argued, Chancellor Merkel was already downplaying the impact of broad counterterrorism-motivated NSA surveillance when further Snowden revelations exposed widespread spying on European leaders. 166 It was at that point that she told the President that “she unmistakably disapproves of and views as completely unacceptable such practices” and that “[s]uch practices have to be halted immediately.” 167 The German Attorney General opened an investigation into the tapping of Chancellor Merkel's phone. 168
But much more is potentially at stake than catching an earful from a foreign leader. The revelations had the effect of recalibrating the relative power of privacy-minded European politicians, who had been successfully sidelined by more security-minded officials for a decade. Questions have been raised about the ongoing vitality of the Terrorist Finance Tracking Program (TFTP) and Passenger Name Records agreement, two central pillars of U.S.-European cooperation in counterterrorism. And the European Court of Justice recently struck down the U.S.-EU Safe Harbor agreement, which, for the past fifteen years, enabled American firms to self-certify that they take certain precautions with respect to personal data. In the aftermath of the Snowden revelations, European officials argued that a national security and criminal justice carve-out had been interpreted too liberally by American officials. The European Court of Justice concluded that national-level officials within the European Union should oversee data privacy, upending the compact, which governs more than 4000 companies. Paul Nemitz, a director in the European Commission's justice department who is overseeing a new data privacy regime for Europe, recently expressed the view that the legal authority that “empowers the NSA to basically grab everything which comes from outside the United States, is a real trade barrier to a European digital company to provide services to Americans inside America.”

Certain countries, like Germany, have pursued--apparently without success--“no spy” agreements with the United States, expressing the desire to be treated on par with the so-called “Five Eyes,” an intelligence alliance that unites the United States, United Kingdom, Australia, New Zealand, and Canada. More dramatically, a number of countries have already taken steps in the direction of imposing data localization requirements. Google Chairman Eric Schmidt recently characterized this development as carrying the potential to “break[] the Internet” because foreign governments are “eventually going to say, we want our own Internet in our country because we want it to work our way, and we don't want the NSA and these other people in it.” Blowback by allies against American intelligence practices has largely focused on electronic surveillance practices. But it has also extended to human intelligence collection. After the arrest of a German intelligence officer alleged to be spying for the United States, Germany decided, for the first time since 1945, to engage in counterespionage against the United States and the United Kingdom. In the meantime, American officials have reportedly engaged in an unprecedented cessation of espionage against European governments.

In sum, the Snowden leaks have galvanized technology firms and allies to join longstanding skeptics of the surveillance state, like privacy groups, in putting pressure on the White House to resist the agenda of the intelligence bureaucracy. To be certain, the market- and strategy-based incentives that motivate these actors are morally shallower, and for that reason more malleable, than the stances taken by the nongovernmental organization critics of surveillance. Zooming out from the issues at hand, it is possible to view the likes of Google and Facebook, with their own insatiable appetites for data, as exhibiting some of the same features--and implicating the same sorts of concerns--as the intelligence agencies against which they rail. And certainly the allied governments that have lodged complaints with the United States, in addition to lacking clean hands themselves, are somewhat parochial in their outlooks. Their central (stated) preoccupation is with the privacy of their own officials and citizens, not those of the United States. But notwithstanding these complications and limitations, with powerful actors engaged in pushing back against American intelligence collection, it is clear that the President, perhaps for the first time, has something to gain-- and a lot to lose--in the oversight of intelligence collection.

B. The Shape of Presidential Intelligence
In response to the Snowden leaks--and their catalytic effect on powerful interest groups--the President has looked to curtail the political damage and to respond to interest group pressures by intervening in the area of surveillance policy. The tools that he has employed can be thought of as largely obeying the logic of “centralization” in Moe’s sense of White House-based leadership. The turn to presidential control can be seen across a number of distinct domains. First, there is the straightforward but noticeable phenomenon of the President becoming seized of the issue. As noted above, in the aftermath of the Snowden leaks, President Obama convened various ad hoc groups staffed by a combination of senior officials and outside experts. With their input, and with the benefit of numerous direct meetings with the leadership of privacy groups, technology firms, and allied governments, the President considered, in a newly systematic fashion, the yawning gaps that had emerged in intelligence governance.

The next--and arguably most significant--aspect of the assertion of President Obama's control in this area was the issuance of a presidential directive, one of the typical vehicles of presidential administration. On January 17, 2014, President Obama issued Presidential Policy Directive 28 (PPD-28), articulating “principles to guide why, whether, when, and how the United States conducts signals intelligence activities for authorized foreign intelligence and counterintelligence purposes.” PPD-28 is divided into four sections: (1) principles governing signals intelligence (SIGINT) collection; (2) limitations on bulk SIGINT collection; (3) alterations to the process for SIGINT collection; and (4) requirements and techniques for safeguarding personal information in the SIGINT collection process and reporting requirements for the intelligence community. It is the directive's third section that speaks most straightforwardly to the formation of a new, White House-driven approach to intelligence oversight. Characterizing that change in his address, the President called for “strengthen[ing] executive branch oversight of our intelligence activities” by ensuring that the White House “will review decisions about intelligence priorities and sensitive targets on an annual basis so that our actions are regularly scrutinized by [the President's] senior national security team.”

The key move here is to define the potential risks associated with intelligence practices (and their possible revelation) broadly. As the President went on to explain, the oversight will take into “account our security requirements, but also our alliances; our trade and investment relationships, including the concerns of American companies; and our commitment to privacy and basic liberties.” This echoes a recognition in PPD-28 that intelligence practices--especially insofar as they become public--potentially entail risk to:

- our relationships with other nations, including the cooperation we receive from other nations on law enforcement, counterterrorism, and other issues; our commercial, economic, and financial interests, including a potential loss of international trust in U.S. firms and the decreased willingness of other nations to participate in international data sharing, privacy, and regulatory regimes ....

Institutionally, the NSC--“the major centralizing institution” in foreign affairs--is the place within the White House where many of these competing equities are put on the table and discussed in comprehensive fashion. As Zegart has persuasively argued, the NSC, which began as a site of coordination for the leading cabinet agencies on matters of national security, has evolved into a highly centralized mechanism for consolidating policymaking within the White House at the expense of those cabinet offices. With a National Security Advisor empowered by Executive Order 12,333 on matters of intelligence, and with a large and influential staff, the NSC is the natural venue for effectively leveraging a variety of expertise from across the national security community (writ large) in shaping intelligence policy. Issues can be considered at the staff level, and elevated from there to ever more senior policymakers (through
Deputies and Principals meetings), ultimately arriving at the President's desk. Moreover, the NSC framework allows for participation by a wide range of institutional actors with diverse perspectives, including the Treasury and Commerce Departments and the U.S. Trade Representative. As PPD-28 puts it, “[i]t is ... essential that national security policymakers consider carefully the value of signals intelligence activities in light of the risks entailed in conducting these activities.” Rather than allow the intelligence agencies themselves to shape intelligence policy, a review process headquartered at the NSC increases the likelihood that surveillance practices will be (or become) “truly presidential by hearing [the] views [of senior national security staff], enlisting their expertise, coordinating their contributions, and directing policy toward presidential ends.”

In overseeing intelligence collection, the President is also able to draw on the distinctive outlooks and expertise of a range of other White House entities, like the NEC and OSTP, as well as of senior advisors not devoted solely to national security matters, such as the White House Counsel and the Chief of Staff. On a number of occasions, President Obama has tapped his Chief of Staff, Denis McDonough (who previously served as Deputy National Security Advisor), to carry out sensitive, intelligence-related missions. For example, after talking directly to Chancellor Merkel about the most recent allegations of American eavesdropping in Germany, the President—in a highly unusual move—dispatched his senior counterterrorism advisor and McDonough to Berlin to engage in bilateral discussions with their counterparts. The meeting’s significance was further displayed when the White House published a “readout” of the meeting on its website, noting that “Mr. McDonough and Mr. Altmaier [Merkel's chief of staff] agreed to set up a Structured Dialogue to address concerns of both sides and establish guiding principles as the basis for continued and future cooperation. The Structured Dialogue will be overseen by the Chiefs of Staff.” More recently, the President sent McDonough on another unconventional mission to Senator Dianne Feinstein's San Francisco home, where his job was to negotiate redactions to the Senate Intelligence Committee report on CIA interrogation practices.

The theme of White House-based policy review of intelligence collection was amplified a few months after the issuance of PPD-28 by the President's top counterterrorism advisor Lisa Monaco, who acknowledged that before Snowden, because the government did not conceive that collection programs and operations would be made public, it did not undertake “cost-benefit analysis” of the foreign policy and economic impacts of American surveillance. She conceded that in the post-Snowden era such assessments are necessary, and that they require the involvement of “senior-level policymakers” as well as a “procedure” for effectuating the assessment of potential tradeoffs. Indeed, by the time she addressed the matter publicly, Monaco had already “overseen weekly interagency task force meetings since August that had included representatives from the Office of the Director of National Intelligence, the Pentagon, and the State Department; cybersecurity experts; economic analysts; and lawyers from the White House Counsel's Office.” The White House also announced a similar process for decisionmaking on the related issue of “Zero-Day” vulnerabilities, accepting the essence of the recommendation of the President's Review Group that vulnerabilities ought to be presumptively patched, and that “[b]efore approving use of the Zero Day rather than patching a vulnerability, there should be a senior-level, interagency approval process that employs a risk management approach.”

III. ASSESSING PRESIDENTIAL INTELLIGENCE

Assessing presidential intelligence requires coming to terms with a number of complex tradeoffs. Presidential intelligence has the capacity to promote more effective, accountable, and rights-regarding intelligence practices. But it also entails a number of potentially significant downsides, including diminished intelligence expertise and enervated congressional oversight. Under certain specifications, it might even foster the conditions for abusive practices.
Presidential administration is experienced with these sorts of tradeoffs and has generated a robust scholarly literature that can help illuminate how these issues might play out in the intelligence setting. In this Part, I draw on that literature to highlight, first, some potential upsides, and then some potential downsides, of presidential intelligence, all the while acknowledging the ways in which the intelligence bureaucracy conforms to its own logic and institutional pressures.

A. The Benefits of Presidential Intelligence

The most significant potential benefits of presidential intelligence are its capacity to promote more effective, more accountable, and more rights-regarding intelligence. The first two of these strengths run parallel to claims that have been advanced on behalf of presidential administration. The third claim sets presidential intelligence apart.

1. Strategically Sound Intelligence.--Presidential intelligence entails a centralized mechanism for reviewing intelligence practices in light of their overall consequences, a job that requires the inputs of policymakers, and so cannot be performed within the intelligence bureaucracy itself. The motivating ideas here are as simple as they are attractive. Intelligence collection practices ought to be assessed for their efficacy and employed only to the extent that their overall benefits exceed their costs. Determining the appropriate scope of intelligence gathering entails calling forth a wide range of perspectives and expertise, a task to which the White House is well suited. The idea that the NSC is able to convene an interagency process through which to arrive at better-calibrated intelligence collection resonates powerfully with a body of academic literature that assesses the role of OIRA in performing centralized review of regulatory decisionmaking. Inaugurated by an executive order issued by President Reagan, but famously retained by all subsequent Presidents, OIRA review institutionalized the White House's role as a clearinghouse for important rulemaking across the government. While OIRA unquestionably arose out of the Reagan Administration's deregulatory agenda, it is a subject of ongoing debate whether that antiregulatory posture is built into the nature of OIRA's work. Some academics maintain that such a bias continues to characterize the office's outlook, while others-- especially veterans of the office--resist that account, maintaining that OIRA is not "necessarily blindly hostile to agencies or to regulation" and that its largely apolitical staff are more "ideologues for efficiency" than antiregulatory in outlook. Another debate revolves around the degree of secrecy that attends OIRA decisionmaking, with critics raising questions about the transparency of its website, its lack of written or public communication, and its opaque review process.

OIRA performs three key functions. The best known is the office's employment of technical cost-benefit analysis (CBA) to review the economic soundness of proposed regulations. But the other two (considerably less heralded) roles that OIRA discharges are, if anything, more directly relevant to the prospects of presidential intelligence. First, scholars and former practitioners have taken stock of OIRA's critical function as an "information aggregator." OIRA asks multiple departments and agencies for their views in order to corral information and expertise and help harmonize regulation across agencies. As former OIRA Administrator Cass Sunstein put it, a defining feature of the office is "the idea of interagency coordination and consultation, so that when a rule [comes] from the U.S. Department of Agriculture, the people at the Department of Justice and the Environmental Protection Agency [are] aware. That's really valuable."

Second, OIRA performs nontechnical, nonmonetized CBA of proposed rules, tallying and comparing their upsides and downsides in conceptual terms. As Professor Amy Sinden has recently argued, the very project of CBA can and should be thought of as containing two key elements that are analytically and practically severable. One is the technical comparison of monetized costs and benefits of proposed government action. Another is the more conceptual tradeoff analysis in which costs and benefits are compared without attempts at rigorous quantification.
In recommending an OIRA-like function for intelligence, I emphasize the desirability of these two defining features—harmonization across various agencies and the application of nontechnical cost-benefit logic—which are ripe for export. Harmonization in the intelligence context entails a mechanism for gathering inputs from all the spy agencies (which differ from one another mainly in terms of the surveillance technology they employ) and analyzing those inputs in light of the more comprehensive strategic outlook that the NSC uniquely possesses. Even Judge Posner, who (in his scholarly work) generally opposes centralization of intelligence functions, has argued that more coordination on intelligence collection modalities is important. Among other things, it will heighten efficiency by nipping in the bud a trend toward unnecessary duplication of effort in collection—for example, the periodic flare-ups between the NSA and CIA in allocating surveillance authorities.

Presidential intelligence can also benefit from nontechnical CBA when considering various potential courses of intelligence collection. This sort of decisionmaking also belongs in the White House as part of an expanded NSC process. As noted above, in a post-Snowden era, the potential strategic and economic costs (and benefits) of intelligence collection are critical inputs for assessing the overall advisability of proposed intelligence collection. As Goldsmith has put it, government “must balance the security benefits of NSA activities against vociferous privacy and legitimacy concerns at home and against significant potential economic fallout for US firms’ global business.” This sort of balancing should not be limited to the consideration of the downstream consequences of electronic surveillance, but should also be applied to other intelligence collection methodologies, including human intelligence. For example, in a recently published monograph, intelligence scholar and current Chairman of the National Intelligence Council Greg Treverton argues that a thorough review of the intelligence value of American human intelligence is warranted, writing that “[f]rom what is publicly known, the record is not impressive.”

In many cases, the costs that must be weighed will show up only in the event the public finds out, in which case the likelihood (or perhaps more realistically, the timing) of revelation must itself be factored into the analysis. This is a straightforward application of what is often referred to as the “Front Page Rule,” in the sense that officials ought to make only those decisions that would withstand scrutiny on the front page of a newspaper. The President’s Review Group included the Front Page Rule in its list of recommendations for surveillance reform, and Goldsmith recently charged intelligence lawyers to take the rule seriously. That said, the track record of intelligence agencies in this area is, at best, lackluster. As Bruce Schneier has written:

While the NSA excels at performing ... cost-benefit analysis at the tactical level, it's far less competent at doing the same thing at the policy level. The organization seems to be good enough at assessing the risk of discovery—for example, if the target of an intelligence-gathering effort discovers that effort—but to have completely ignored the risks of those efforts becoming front-page news.

Even a cursory comparison with existing intelligence oversight institutions reveals the relative advantages of the White House-- and in particular, the NSC--in assessing these sorts of risks. It goes without saying that neither the FISC nor courts of general jurisdiction have anything like the requisite capacity to harmonize disparate voices from across the intelligence community or to promote a set of intelligence policies that balance competing strategic imperatives. Congressional committees are more capable on this dimension than courts, for their remit includes ‘ensuring that taxpayers’ funds are spent appropriately and efficiently on programs and activities that produce useable intelligence information [and] that intelligence activities are effective in protecting the United States and its interests from foreign
threats.” Nevertheless, congressional overseers lack real-time access to the full range of inputs necessary for integrating intelligence collection practices into a national strategy. Furthermore, the fragmentary nature of congressional oversight committee jurisdiction impedes the development of a comprehensive outlook. Meanwhile, legalist checks within the executive branch, ranging from inspectors general to offices of general counsel, are not designed or staffed, except obliquely, to analyze questions of intelligence efficacy. Consider in this regard the decision (exposed in one of Edward Snowden’s leaked documents) to tap German Chancellor Angela Merkel’s personal cellphone as far back as 2002 (when she served as leader of the opposition in the Bundestag), a decision seemingly made by the NSA without the President’s knowledge. Whatever legal barriers that may have applied to this collection were dwarfed by strategic and economic concerns of the sort that legalist oversight bodies are not called upon to address. It is precisely in considering these downstream policy-based consequences that presidential intelligence would have proved useful. Nested at the hub of the national security state (understood in the broadest sense to contemplate instruments of military, diplomatic, and economic power), White House review would have brought a holistic approach to calibrating the proper metes and bounds of this and other collection efforts, recognizing that how intelligence is gathered must be integrated into any national strategy.

As noted above, the OIRA analogy should not be stretched too far. It is difficult to see how technical CBA could be performed in this area. As Professor John Coates recently cautioned with respect to the extension of technical CBA to financial regulation, any attempt to perform such an analysis of intelligence collection is likely to supply an occasion for camouflaging qualitative judgments, rather than getting hard-edged analytic traction on the issues. But to conclude that intelligence oversight “does not take place in a political vacuum in which legislators conduct a Spock-like assessment of options, costs, and benefits” is not to diminish the possibility and desirability of a more conceptual review for tradeoffs carried out by the White House. And the absence of mathematical rigor does not absolve overseers of responsibility for generating a meaningful methodology to assess the efficacy of intelligence. As noted above, the absence of clearly defined metrics has been a major shortcoming of intelligence oversight (and the oversight of national security policies more generally), playing into the hands of a more ideologically inflected and less pragmatic discourse about intelligence. As one PCLOB member has observed, “the natural tendency of the government, the media, and the public is to ask whether a particular program has allowed officials to thwart terrorist attacks or save identifiable lives.” But “plots averted” is hardly the right barometer of intelligence success in counterterrorism, putting aside the obvious problem that this metric does not translate to non-terrorism-related intelligence work. In discussing the 215 metadata program, former NSA Deputy Director John C. “Chris” Inglis, acknowledged that the program was “not a silver bullet in and of itself,” but maintained that its true value inhered in plugging a potential intelligence gap “that we don’t know any other way to cover.” It may be that reaching meaningful conclusions about the value of intelligence collection efforts necessitates a more conceptually rich vocabulary for capturing and measuring intelligence efficacy than we currently have. But from a practical point of view, it is certainly not impossible to assess the utility of information learned from a particular program, source, or collection method.

The desire to push the PCLOB toward considering the policy implications of surveillance practices is commendable. Although its mandate did not make it inevitable that the Board would define its role in largely legalist terms, the fact that the PCLOB is staffed exclusively by lawyers may have pointed it in that direction. The Board’s emphasis on determining, in a court-like fashion, the legality of various programs has arguably detracted from its ability to pronounce more holistically on the programs’ costs and benefits. That said, proposing “that the NSA and other members of the Intelligence Community develop metrics for assessing the efficacy and value of intelligence programs, particularly in relation to other tools and programs,” misunderstands the nature of the measurement that needs to take place.
Members of the policy community, as consumers of intelligence (informed, to be sure, by the insights of intelligence officials), will need to assess the efficacy of intelligence. Intelligence officials suffer from tunnel vision and lack the strategic horizons (and perhaps also the neutrality) to evaluate their own work.  

None of this is to say that the intelligence community can be expected to cede turf willingly to the White House. To the contrary, the capacity of intelligence agencies to engage in what Professor Jennifer Nou, in the context of agency stonewalling of OIRA, has labeled “self-insulating” behavior, is formidable. One way to resist greater oversight is for agency officials to highlight the scope of current (legalist) mechanisms. This sort of approach is audible in the way that intelligence officials talk about legalist oversight. When former CIA General Counsel Stephen Preston speaks of intelligence as a regulated industry and when former NSA General Counsel Rajesh De refers to the section 215 metadata program run by the NSA and FBI as “one of the most highly regulated programs in the federal government today,” they are reading from a common legalist script. Strictly speaking, these officials are not wrong; there is significant legal regulation of at least certain aspects of American intelligence collection. But the presence of this kind of oversight-- and the willingness of officials to tout it--does not speak to the enormous policy discretion that spy agencies otherwise enjoy with respect to intelligence gathering.

In their efforts at self-insulation, intelligence officials are enabled by a climate of secrecy, which limits scrutiny from outside, or even from within, the national security state. Indeed, leaders of the spy agencies themselves struggle to account for all that the organizations they run do. Referring to unclassified Special Access Programs (SAPs), then-Undersecretary of Defense for Intelligence (and current DNI) James R. Clapper observed: “There's only one entity in the entire universe that has visibility on all SAPs--that's God.” The ability of intelligence agencies to dampen White House control is also underwritten by their recognition that the President is ultimately dependent on the work they do, and will, accordingly, only bear down on them so much. As Professor David Cole recently observed, reflecting on CIA Director John Brennan's nuanced response to the recently issued Senate Intelligence Committee's study of the CIA's detention and interrogation program:

“...Both are personally opposed to what went on and deeply troubled by what went on and agree that it should never happen again. And both are ultimately dependent on the C.I.A. for important national security services.”

But in the end the right question is not whether the White House will struggle to extract information from the intelligence agencies, but instead whether the White House can be expected to perform more effectively in this regard than other oversight institutions. Whether the comparison is to congressional committees, the FISC, or the PCLOB, the answer is yes. As noted above, the President and his staff have daily contact with senior leaders of the intelligence apparatus and have at their disposal mechanisms to extract information, up to and including the ability to replace agency heads. Furthermore, and more subtly, the President is positioned to exploit rivalries across the intelligence and national security bureaucracy to defeat the self-insulating strategies that individual agencies might pursue.

2. Accountable Intelligence.--Under many (though, as discussed below, not all) specifications, presidential intelligence carries the potential for heightened democratic accountability. In one sense, the logic here verges on the tautological. Substituting presidential intelligence for a system that historically empowered the permanent intelligence bureaucracy to self-regulate promotes responsiveness to the people's elected representative. As one intelligence scholar has explained, “what some may perceive as a president's 'preconceptions' and 'biases,' may well be the entirely proper policy orientation that a president was elected to pursue.” But upon closer inspection, more nuanced judgments can be teased out, and two distinct concepts of accountability come into view. First, there is the way in which presidential intelligence underwrites (and is underwritten by) what might be thought of as a pluralist account of accountability. As noted above,
the President has repeatedly interacted with emergent interest groups in the intelligence domain, especially foreign heads
of state and diplomats on the one hand, and technology and telecommunications executives on the other. Precisely
because these conversations are relatively intimate and entail discussions with knowledgeable intelligence insiders (and
efforts by officials to mollify actors they need to keep on board), they are likely to involve candid talk about intelligence
practices. In turn, these candid exchanges can be said to supply a measure of accountability with the interest groups
standing in for (at least some portion of) the general public.

*685 Second, presidential intelligence has inched toward a more straightforwardly democratic vision of accountability,
a turn which is itself dependent on the heightened visibility of the intelligence apparatus. The allied presidential
administration literature regards the President's ability to “go public” as a defining feature of that project.265 The
President and his senior advisors have spoken publicly and extensively about the changes at hand. For example,
in a major speech that he delivered to accompany the issuance of PPD-28, President Obama sought to reassure a
skittish public that the United States grapples meaningfully with the political, ethical, and legal dilemmas posed by
contemporary surveillance.266 As the President put it, “we will reform programs and procedures in place to provide
greater transparency to our surveillance activities.”267 Other top officials have made extensive public appearances,
perhaps none more so than the incumbent ODNI General Counsel, who has spoken repeatedly to specialist and
nonspecialist audiences.268 This turn to greater openness is consistent with Professor John Ferejohn's insight that
heightened visibility is causally linked to greater official power.269 To take an example from the intelligence domain,
the PSP, attended as it was by intense secrecy, ultimately could not bear its own political weight; when Congress openly
legislated intelligence gathering comparable in scale to the PSP but more visible to the public, the program could pass
muster.270

*686 The presidential turn to greater publicity is itself made possible by his ability to reach out to a national (and
perhaps just as importantly, global) audience.271 A key component of that power is the President's unilateral capacity to
declassify information and to reveal aspects of intelligence programs.272 By contrast, neither the FISC nor congressional
intelligence committees may declassify information on their own authority,273 which severely constrains the ability of
those institutions to engage the public directly on intelligence matters. This power has been powerfully on display in the
post-Snowden environment. Working with the DNI,274 the White House has acted to make public many previously
classified documents, including the intelligence community's annual priorities, countless surveillance-related documents,
and a series of important opinions of the FISC.275 That is not to suggest that the White House possesses a monopoly
on official discussion of intelligence matters. Some scholars have drawn attention to the capacity of the congressional
oversight committees to “explain[] and represent[] the intelligence community to the public,”276 and certain legislators
have, in fact, taken to the airwaves to comment (critically277 and approvingly278) on pressing intelligence controversies.
The recently issued Senate *687 Intelligence Committee report on CIA detention and interrogation279 attests to the
ongoing power of Congress to speak directly to the American public. But the White House will always have an edge in
terms of its access to the most timely information and the relevant strategic context.

I hasten to add that gains here are to be understood in comparison with the prior state of affairs. For a host of reasons,
democratic accountability cannot be fully realized in the intelligence area (if it can be anywhere). First, notwithstanding
the discussion above regarding the tight connection between presidential intelligence and heightened publicity, the
default setting for intelligence programs and oversight remains secrecy. All presidential revelation in this area, however
well motivated and ostensibly thorough, is necessarily incomplete. Furthermore, greater visibility is not the same as
transparency. That is because the power to disclose selectively is parasitic on, and marbled into, the power to conceal.
Paraphrasing Senator Daniel Moynihan's classic work, one might say that selective disclosure, like secrecy, is itself a “form of regulation.” In the fullest sense the President cannot be said to be securing the assent of the people for policies that are known to the people only through selective disclosures made by the White House. As noted above, the pluralist accountability sustained by presidential interaction with informed interest groups may be more promising on this dimension. Second, as Professor Jide Nzelibe has argued, the very idea of a nationally elected President capable of conferring democratic legitimacy on White House decisions rests, at least in part, on myth. By and large, national elections do not pivot on the politics of this or that issue --certainly not surveillance.

But these qualifiers themselves need qualifying. Allowing that public knowledge of, and participation in, intelligence governance is constrained by secrecy, there is nevertheless a great deal more information available in the public domain than there ever has been. As compared to the administrative state, where centralized oversight bodies like OIRA are typically more, not less, secretive (and less accountable) than the agencies they superintend, the opposite is true in the intelligence arena. And while it is unlikely to top the list of politically salient issues in the current presidential-election cycle, intelligence oversight can be expected to influence campaign fundraising in certain key sectors of the economy and slices of the voting public.

3. Rights-Regarding Intelligence.--Presidential intelligence may well mean more privacy-oriented intelligence, as compared with the baseline. This is true at the conceptual level. Intelligence collection that is better aligned with strategic judgment is more likely to pass muster under the Fourth Amendment, according to which reasonableness is a touchstone for establishing legality. But it is also true in a more operational sense. Greater political control from a White House under economic and strategic pressure from technology firms and allies may also yield more privacy-oriented intelligence. For example, as noted above, PPD-28 embodies a commitment to extend certain privacy protections to non-U.S. persons. This ratcheting up of privacy protections beyond the dictates of any statute or the Fourth Amendment--“a major change in U.S. policy”--dovetails with the interests of allies and global firms seeking to reassure skittish citizens and customers. In the PPD's demands that signals intelligence “be as tailored as feasible” and that bulk data not be used for affirmative foreign intelligence gathering, the pressure from technology firms and allies is also detectable.

ACLU lawyer Ben Wizner's observation that “one of the great contributions that Snowden has made is to make some very powerful tech companies adverse to governments” captures something true about the emerging dynamic. Technology firms, channeling global consumer demand, are currently serving as a catalyst for more constrained surveillance practices. It is precisely the conditions of privatization and mutual dependency between the technology firms (and allies) and the government that have underwritten the power of these actors to push back. The Madisonian insight that individual rights are most effectively protected when “[a]mbition ... [i]s made to counteract ambition” --a claim that is usually realized through inter- and intragovernmental checks at the federal and state levels--is here operationalized when technology firms and allies, looking to secure their own interests, pressure the White House to push back against the intelligence community.

Presidential administrations may, of course, be more or less receptive to what the technology firms are saying, but given these firms' ever-growing clout within the American economy, it is hard to imagine a President ignoring their demands altogether. Obviously a shift in the threat environment could have a powerful impact. But the claim that presidential intelligence is uniquely vulnerable to dynamic assessments of risk misses the mark, because all intelligence oversight institutions (courts, congressional committees) have tended to buckle under the pressure of a current or very recent
national security emergency. That said, one potential worry about the capacity of presidential intelligence to supply rights protections stands out: the mounting threat of cyberattacks. Because of the technological interdependence of surveillance and cybersecurity, certain privacy protections that have largely been obtained in the context of pushing back against counterterrorism surveillance programs may come under increasing pressure from the mounting concern over cybersecurity.

Championing the capacity of political control to underwrite rights-protection is not to gainsay the value of other mechanisms for promoting rights-regarding intelligence. All three branches of government have important roles to play. Indeed, as I discuss below, presidential intelligence might itself enhance the capacity of other overseers to protect privacy. For example, presidential control might induce more sensitivity to law and legal controls within intelligence agencies by shining a light (if only because more officials, including in the White House, will be scrutinizing the work of intelligence) on potentially shaky legal theories. For that matter, the prospect that the White House might be paying attention could induce intelligence lawyers to be more scrupulous in making representations to the FISC and in adhering to that court's mandates. But the central point remains that presidential control, predicated, at least initially, on interest group participation, is likely to be a potent vector for meaningful (and enforceable) rights-regarding reforms.

Here, too, the benefit of presidential intelligence needs to be understood in relation to the preexisting baseline. Although the Snowden leaks showcase an intelligence bureaucracy that has largely internalized its responsibility to make a good faith effort to obey the law, and reveal effectively no officially sanctioned abuse, the leaks also show how existing institutions have fallen short in ensuring the legality of intelligence collection. Here the well-known story of domestic metadata collection on the (claimed) authority of section 215 of the Patriot Act is instructive. The government program rested on a legal interpretation that made it through a gauntlet of institutions designed to check for legally questionable intelligence gathering, even though doubts surfaced about the program's legality (and efficacy) at a number of points along the way. It was briefed to members of Congress, countenanced by the OLC, and passed on by the FISC, which, after some back and forth, permitted the government to extend out from the initial target “three hops”-- meaning that the government could look into a pool of metadata massively larger than that belonging to the individuals being investigated. Until Congress passed the USA Freedom Act of 2015, prohibiting the bulk collection of all records under section 215 and mandating that the government base applications for this kind of data on a “specific selection term,” it was the President's engagement with the issue, inspired by the feedback he received from American technology and telecommunications firms, that had produced the most movement to cut back on this authority and to interpose a measure of accountability.

The exception of the recent legislation proves a larger rule. Congress, under conditions of seemingly unprecedented partisan rancor, is unlikely to reset the basic terms of the bargain between the American people and the intelligence apparatus. Meanwhile, neither beefed-up civil liberties offices within the various agencies nor investigatory bodies without remedial authority like the PCLOB possess the institutional heft on their own to enact or enforce serious reforms. And although, as noted below, courts are increasingly involved in shaping intelligence policy, presidential intelligence is more nimble and in some sense more ambitious in what it can achieve (at least in the short run), for example when it comes to extending privacy protections beyond America's borders.

B. Three Potential Downsides

Although it is potentially conducive to more effective, accountable, and rights-protective intelligence, presidential control entails certain risks. I discuss three in particular: interfering with expertise, fanning the flames of partisanship, and
threatening abuse. I regard the first two concerns as essentially surmountable, or at least no more damaging to the case for presidential intelligence than comparable worries that surface in connection with presidential administration. The third concern is unique to the intelligence environment and necessitates thinking that is attuned to the dispiriting history at hand and alert to potential ways to prevent it from being repeated.

1. Politicization. -- Striking the balance between political control and agency expertise is a core tension that runs throughout the administrative state. Agencies in a sense owe their existence to a claim of technical know-how that they are able to deploy in the service of sound policymaking. But that commitment to expertise trades off against competing aspirations to democratic accountability rooted in the close ties between the agencies and their political overseers. The presidential administration literature is attuned to this dilemma. In her 2001 article, then-Professor Kagan acknowledged that “an important place for substantive expertise remains in generating sound regulatory decisions” and that “[t]o the extent that presidential administration displaces this feature of agency decisionmaking in areas where it legitimately should operate, this substitution effect must weigh against the practice.” But cordonning off science from politics is famously knotty even in the abstract. And the problems do not get easier when political actors and institutions are engaged. A recent example of the politicization of agency expertise came from President George W. Bush's EPA and its skepticism of climate science. The episode culminated in a rebuke from the Supreme Court motivated, in Professors Jody Freeman and Adrian Vermeule's telling, by “the Court majority's increasing worries about the politicization of administrative expertise.”

In the intelligence setting, a similar set of issues has played out for generations, concerning the well-known phenomenon of politicization of intelligence. Indeed, the very worry about politicization is “a US invention, one stemming from the specific structure and role of the US intelligence community in the recurring struggles over strategic issues in defense and foreign policy during the Cold War.” The debate pits two schools of thought against each other. On the one hand, there is the view, associated with intelligence scholar Sherman Kent, that intelligence and policy must remain separate. On the other hand, there is a view that is associated with former-CIA head Robert Gates, which insists that too much separation impedes the purpose behind intelligence: to generate useful and relevant insights for policymakers. The balance within the intelligence state has historically tipped toward the Kent view. For example, Paul Pillar, with the Iraq weapons of mass destruction (WMD) fiasco in mind, offers a powerful defense of the value of resisting political pressure in shaping intelligence estimates. Scholars like Jennifer Sims, however, have pushed back: U.S. intelligence officers often do not seem to believe they are working on behalf of policy makers or as part of their team. They tend to see themselves as a check on an administration's power and the repository of truth in a system riddled with biases .... Although policy makers do want intelligence to provide facts or “ground-truth,” other branches of government have the job of checking the power of those in office, not intelligence.

Corresponding to this view, and tending to reinforce it, is a nascent trend to comingle intelligence collection and intelligence analysis. This is another formerly ironclad barrier that has begun to give out under mounting evidence that producing the most valuable intelligence requires persistent interaction between operators and analysts. This new idea is based on an old one, part of a road not taken when a central intelligence agency was chosen over an alternative institutional design that would have co-located all intelligence analysis with various government departments
responsible for policymaking. Here, too, critics invoke “politicization” as a would-be knock-down argument in favor of maintaining, or even fortifying, the traditional divide.

Regardless of how the balance is struck in terms of the roles of policy and politics in shaping substantive intelligence judgments, the issue of politicization takes on a somewhat different cast when it comes to heightened presidential oversight of intelligence collection modalities. It makes sense to quarantine from politics the factual inquiry into whether Saddam Hussein possessed WMD, or how far along the Iranian government is in acquiring weapons-grade nuclear material. But it is not comparably intuitive—and in fact, makes little sense—to bar the White House from expressing a view on the desirability of spying on an ally, or weighing in on whether to forego controversial programs like metadata collection because their costs may outweigh their benefits. The intelligence agencies have no claim to comparative advantage here. Indeed, concerning the overall assessment of the value of intelligence programs in relation to overarching goals of strategic and economic statecraft, the spy agencies are likely to be less informed (even cumulatively) than the White House, which, as discussed above, can summon the perspectives of multiple “customer” agencies to develop a comprehensive picture. Furthermore, because these sorts of judgments inevitably traverse the fact-value divide, they are appropriate for White House decisionmaking. As Kagan puts it, “[a]gencies ... often must confront the question, which science alone cannot answer, of how to make determinate judgments regarding the protection of health and safety in the face both of scientific uncertainty and competing public interests. With respect to these matters, a strong presidential role is appropriate ....”

But that doesn't fully respond to the potential shortcomings of presidential intelligence in this regard. First, although stylized presentations of the intelligence process sharply distinguish between collection and analysis, as discussed above, that boundary is increasingly murky. As a result, presidential intelligence might tend to tip the balance on analytic judgments in favor of the White House's preferred views, especially concerning relatively low-profile issues. Second, as Zegart has observed, the CIA, subject to political pressures imposed by the White House, has veered away from its founding mission of forecasting strategic threats, and has taken on a much more tactical focus. Presidential control might “politicize” intelligence in the sense of further imposing a short-term outlook on what are meant to be more long-term strategic intelligence estimates. Third, politicization might take place in the White House itself, to the extent that the President's political advisors, mindful of election dynamics, encroach on the more policy-oriented decisionmaking by the NSC.

Recognizing the intractability of these issues, it is worth recalling that presidential intelligence is not a unique vector for these sorts of pathologies. In his capacity as intelligence Consumer-in-Chief, the President arguably has more ability to shape (and distort) intelligence priorities than in his role as strategic overseer of collection. But these tensions are clearly real and, over the long term, could threaten the integrity of key aspects of the intelligence enterprise. Intelligence scholar Professor Joshua Rovner argues that the only antidote to politicization is a return to greater secrecy. Even setting aside the normative implications of that claim, such a plan is practically unattainable for reasons discussed above. Instead, good institutional design— for example, replication of aspects of the “finding” process (as discussed below)—represents a more promising avenue for addressing or preempting these concerns.

2. Partisanship.—A different concern focuses on the perils of heightening presidential power at the expense of Congress—and, more generally, of endorsing political controls in an age in which norms of hyperpartisanship have become pervasive across government, up to and including the national security state. Here, too, the normative debates sparked and informed by the presidential administration literature are suggestive. As to both accountability and effectiveness, the presidential administration literature points to certain advantages rooted in the President's status as the nationally
elected leader, as well as the relative shortcomings of Congress, including its limited institutional attention span, the nonrepresentative nature of committees, the committees' state of being captured and their stovepiped regulatory purviews, and the legislature's resource and expertise gaps—all of which get to the heart of why Congress chooses to delegate policymaking to agencies in the first instance. Furthermore, unlike the President, who is incentivized to take ownership of issues because he is held accountable for them regardless, Capitol Hill does not operate on such political logic.

Supporters of congressional oversight meanwhile resist these assumptions. They point to Congress's unmistakable power to control agencies through appropriations and, of course, through substantive legislation. Professor Thomas Sargentich has cautioned that skepticism of the capacity of congressional committees to underwrite democratic accountability should not “defeat the claim that when Congress acts as a whole, with majorities of both the House of Representatives and the Senate in agreement, it represents a broad range of interests, geographical areas, and political orientations.”

Although the academic literature on presidential control of intelligence collection is relatively scant, a rich body of commentary diagnoses the limitations of congressional oversight. In its earlier days, there was a hopeful air about the project. Former CIA Director Robert Gates went so far as to suggest that, starting in 1975, the “CIA ... move[d] from its exclusive relationship with the President to a position roughly equidistant between the Congress and the President.” Furthermore, congressional oversight—especially of covert action—paid dividends early on, including helping to avoid operations that would have produced more harm than good. Gates has observed that “some awfully crazy schemes might well have been approved had everyone present not known and expected hard questions, debate, and criticism from the Hill.”

But over time the limits of congressional oversight came to the fore. For a host of reasons, congressional oversight began to decline, or at least so the familiar story goes. With committee members largely unable to take public credit for their work, investment of time and effort on intelligence oversight defied basic political logic. The problem was only compounded by the difficulty of sharing highly classified information with committee members and congressional staff. Add to that the fragmentation of oversight responsibilities among multiple committees, the lack of meaningful budgeting authority on the part of congressional intelligence overseers, and the limited capacity to deploy police patrol-type oversight, and the limits of congressional oversight come into sharp relief. As Representative Norman Mineta, who served on the House Permanent Select Committee on Intelligence in the Reagan years, caustically observed, “We are like mushrooms .... They keep us in the dark and feed us a lot of manure.” Zegart quotes a frustrated congressional staffer to the effect that “[t]he silver lining with the FBI is that at least they're nonpartisan in their non-cooperation with Congress.” Nor have things improved in the years since Zegart undertook her study. If anything, relations between the CIA and Congress have recently been described as “more fraught than at any point in the past decade.”

And yet other commentators resist this (by now familiar) narrative. Intelligence scholar L. Britt Snider has offered the view that whatever the shortcomings of the current system, “compared with the level of congressional awareness that existed in 1975, the difference is like night and day.” Furthermore, and not trivially, although Congress initially regulated the intelligence community with an exceedingly light touch—the CIA's organic law is breathtakingly short on detail, while the FBI lacks a basic legislative charter altogether—the last decades have witnessed greater congressional
regulation. The initial FISA law *700 of 1978, the Patriot Act, the Intelligence Reform Act of 2004, 348 the FISA Amendments Act of 2008, 349 and the recently passed USA Freedom Act of 2015 attest to this evolution.

Against this backdrop, it is difficult to offer a confident assessment of how heightened presidential intelligence might interact with congressional controls. On one level, there are some reasons to be hopeful that the combination of presidential and congressional control might lead to better overall oversight. For example, as discussed below, any presidential “finding” on sensitive collection programs could then be briefed on the Hill (much as in the parallel case of covert action), teeing up and focusing congressional oversight of collection platforms or modalities. It is suggestive in this regard that the Senate Intelligence Committee has recently evinced an appetite for an enlarged role in overseeing intelligence collection under Executive Order 12,333--including an unprecedented step by that body to catalogue and account for the full spectrum of intelligence gathering. 350 Furthermore, an expanded list of intelligence posts that require Senate confirmation (along the lines I propose below) would increase the opportunities for congressional buy-in. Thinking more structurally, the same interest group pressures that have catalyzed and shaped the exercise of presidential controls are also in play on Capitol Hill, at least as far as privacy activists and technology firms are concerned (the lobbying efforts of allies are less visible in Congress). The recent passage of the USA Freedom Act may imply more sustained congressional attention to issues of surveillance in a way that is likely to be mutually compatible with presidential controls. And at the theoretical level, a leading piece of empirical research into wartime interbranch relations recently concluded that “members of Congress will enact policies that more closely reflect presidential preferences when they assign greater importance to the national vis-à-vis local implications of public policy,” 351 implying that on questions of intelligence policy, Congress and the White House might be fairly well aligned. In sum, it is not at all clear that presidential intelligence will have the effect of “crowding out” or (further) marginalizing congressional intelligence oversight; it is even possible that congressional oversight will be improved.

But these more optimistic assessments might well be eclipsed by the dynamics of hyperpartisanship. Under conditions of “separation of parties,” 352 presidential intelligence can be expected to elicit partisan blowback on Capitol Hill more so than a set of policies identified with the intelligence bureaucracy (rather than the White House) might. Professors Daryl Levinson and Richard Pildes have taken issue with Kagan's optimism about the complementarity of presidential and congressional controls under conditions of divided government, noting that “[w]e should expect that the same party competition under divided government that gridlocks the legislative process and motivates presidential administration will create an adversarial ‘oversight arms race’ between the President and Congress over the bureaucracy.” 353 As a corollary, under conditions of unified government, presidential intelligence might foster a climate in which congressional overseers-- especially those who belong to the President's party--pull their punches. 354

It may be that certain structural features of the intelligence world better insulate it from the partisan dynamics that dominate the work of Congress on regulatory matters, or even the rest of the national security state. 355 For one, some of the most pressing intelligence controversies defy the familiar right-left politics that typically organize debate and inflame partisan sentiment in Washington. 356 Second, some of the background conditions under which intelligence committees work--conditions that, as noted above, tend to interfere with their ability to perform effective oversight--actually insulate them from some of the hard edges of hyperpartisanship. Committees that operate with lean staffing and in secret, and that oversee complex policies and bureaucracies, supply relatively inhospitable environments for partisan grandstanding. That said, as presidential intelligence takes root, and as the lobbying dollars of technology and telecommunications firms reshape the economic incentives of congressional participation, this once *702 sleepy outpost of congressional oversight may take on more of the familiar excesses of partisan politics.
In this light, it remains to be seen how to interpret the dynamics behind the report that the Democratic Chairman and the (exclusively) Democratic membership of the Senate Intelligence Committee issued on CIA interrogations. On the one hand, it can be viewed as a refreshingly nonpartisan effort in which a congressional leader took aim at the CIA despite the fact that she shares a party affiliation with the sitting President, whose close ally runs the CIA. On the other, it can be seen as evidence of heightened partisanship in congressional oversight, in view of the fact that the report principally scrutinizes the CIA's conduct during the Bush Administration and that the Republican members of the Senate committee (and their staff) did not participate in the research or the issuance of the report.

3. Abuse. -- I have argued above that presidential intelligence may well serve to enhance privacy protections. Yet, under certain specifications, fusing presidential power with intelligence capabilities might have the opposite effect, possibly even enabling the sorts of abusive practices that occasioned the significant intelligence reforms of the 1970s. Given the “very extensive history of intelligence activities infringing on the rights of Americans,” this concern is undoubtedly serious. The case of the PSP is instructive. Through the PSP, the White House invoked the President's commander-in-chief authority to authorize certain kinds of intelligence collection that would otherwise have been governed by FISA. Former Congresswoman Jane Harman recently revealed that, as a member of the “Gang of Eight” legislators initially briefed by the White House about the PSP, she was advised that the program was in full compliance with the law. What she did not know at the time, and learned only when The New York Times exposed the program years later, was that the White House's claim to legality rested on a theory that the President is empowered to override statutory law in the area of national security.

To argue (as I have done above) that the PSP was hardly an instance of presidential intelligence as I define it is not to put to rest the more general concern that enhancing the White House's role in spying carries risk for abuse. That said, there are a number of structural factors that tend to reduce the likelihood that notorious episodes such as this will be repeated.

First, as noted above, presidential intelligence presupposes and entails greater visibility of the intelligence apparatus than has ever been the case. To be certain, as discussed above, visibility is not the same as transparency. But--and this is the nub for present purposes--outright abuse is less likely to go unnoticed under conditions of greater visibility, including within the government. The emergence of what Goldsmith refers to as the “synoptic” presidency--the President's state of being pervasively monitored by a vigilant press and civil liberties bar, as well as by internal watchdogs--marks a significant difference between now and the era of abuses that led up to the Church Committee's damning inquest (and even between now and the years immediately after 9/11, when the PSP debuted). Second, and relatedly, as presidential intelligence becomes a matter of institutional habit within the White House, it will become increasingly difficult to operate outside of the internal processes that define it. It is instructive in this regard that the legal and political pushback to aspects of the PSP from within the administration was as robust as it was. Then-Deputy Attorney General Comey and OLC head Goldsmith concluded that one of the PSP programs having to do with internet metadata was illegal, and the issue culminated in a now-famous showdown between the White House chief of staff and senior legal and law enforcement officials who were all gathered at the hospital bedside of the then-Attorney General John Ashcroft. Third, the sheer scale of the contemporary intelligence state (including the number of private actors who are part of its workforce)--coupled with the interest group politics that have coalesced around these issues--also contributes to the unlikelihood that presidential intelligence could bring about a situation in which the intelligence arm runs amok. After Snowden, no President can reasonably count on the obeisance of the intelligence bureaucracy--which includes legions of young techies who may well be inclined to leak any evidence of abusive behavior.
Another reason to believe that presidential intelligence is not likely to translate readily to abuse is the growing availability of judicial review of intelligence programs. At first blush, this is an area where the differences between the administrative and the intelligence states stand out. In the regulatory arena, judicial review is axiomatic, and it typically implicates a judgment by a court about the essential plausibility of an agency interpretation of a statute or piece of regulation. In the intelligence domain, the picture has been quite different. To be sure, as Professors Samuel Issacharoff and Richard Pildes have shown, it is a mistake to assume that national security policies have been or are off-limits to judicial review. But when it comes to the intelligence milieu, judicial participation really has been largely missing, at least until recently. Granted, some amount of judicial review takes place before the FISC, including a mechanism for that court to pass on the basic validity of certain kinds of collection under the FISA Amendments Act (a process that I have elsewhere analogized to hard look review). But quite unlike rationality review of a proposed administrative rule, that process typically takes place ex parte and in secret--a far cry from the framework, established by the APA, through which private litigants challenge administrative action. And in any event, as noted above, the vast majority of electronic surveillance is undertaken pursuant to legal authorities, such as Executive Order 12,333, which do not entail any judicial oversight at all. The same is true of human intelligence gathering, both overseas and domestically. Even where there have been plausible bases for judicial review of intelligence practices in courts of general jurisdiction, judges (at least until recently) have been exceedingly reluctant to weigh in on the merits, interposing familiar doctrinal barriers such as standing and state secrets. But there are some signs pointing to the prospect of more vigorous judicial review of intelligence practices. It is noteworthy that the Second Circuit recently rejected the government's statutory interpretation of section 215, though it withheld judgment on the underlying constitutional issues. And the legality of collection under section 702 of the FISA Amendments Act, an issue that the Supreme Court deflected on standing grounds only two years ago, is now wending its way through the federal courts, on a trajectory likely to culminate in a grant of certiorari. Furthermore, the Supreme Court's recent pronouncements on the intersection of technology and the Fourth Amendment in criminal cases, though ostensibly having no bearing on national security-motivated surveillance, have already begun to inform debate and policy in that area. These developments are themselves linked to the public choice dynamics analyzed above. While privacy groups and activists have historically been responsible for legal challenges to surveillance authority, including, increasingly, before foreign and international tribunals, technology firms are now also involved in a number of high-profile lawsuits that are being contested in the shadow of the post-Snowden recalibration on intelligence. This trend toward increasingly robust judicial checks on intelligence may help to deter and curtail certain potential excesses latent in a cozier relationship between the White House and the spy agencies. Certainly much work remains to be done. As courts tiptoe in the direction of substantive engagement with intelligence issues, they will need to address doctrinal uncertainty on a range of issues, from the questionable ongoing applicability of the third-party doctrine to the scope of extraterritorial application of constitutional and treaty-based privacy rights, to the viability of a foreign-intelligence exception to the Fourth Amendment. It may be that the Supreme Court will eventually have to take its Fourth Amendment doctrine apart and put it back together again. But the fact that that prospect seems increasingly realistic should go some way toward alleviating the worry that presidential intelligence will lead to abusive practices.

IV. DESIGNING PRESIDENTIAL INTELLIGENCE
As documented above, presidential intelligence is in its bureaucratic infancy, and currently operates in a somewhat ad hoc manner. If its potential is to be fully realized, it must be well designed and thoughtfully nested in the larger intelligence-oversight and national security-policy ecosystems. Presidential intelligence ought to be able to leverage the preexisting oversight infrastructure within--and outside--the executive branch. Meanwhile, innovations in presidential intelligence should benefit congressional and other overseers as well. I have two concrete recommendations in mind to help realize these goals. First, as to certain politically, economically, or strategically sensitive or high-stakes collection efforts, the White House ought to be required to involve itself directly in the alignment of intelligence gathering with American interests and values through a process resembling the one employed for covert action. Second, in order to augment accountability within the agencies and to unplug the information flow between the intelligence community and the White House, the ranks of presidential nominees within the intelligence bureaucracy (with and without Senate confirmation) ought to be substantially increased.

A. Presidential “Findings” for Intelligence Collection

The President ought to employ the equivalent of a “covert-action finding”--arguably one of the most successful pieces of oversight architecture from the last generation—to authorize specific, highly sensitive intelligence collection programs. The goal of the finding is to specify and reduce to writing the objectives of the proposed action and to detail the government agencies and any third parties that will be involved. The finding, which contains a nonquantified CBA, is typically reviewed on an interagency basis and then must be personally approved by the President so as to disallow him resort to deniability. The finding must then be submitted to the congressional intelligence committees in a timely fashion. Furthermore, significant changes that come about with respect to covert action once it has been commenced must be captured in a Memorandum of Notification.

The oversight mechanism for covert action came about in three legislative stages, beginning with the passage of the Hughes-Ryan Act of 1974, which required the President to determine that any proposed covert action “is important to the national security of the United States” and mandated reporting to various congressional committees. Next, the Intelligence Oversight Act of 1980 established the relatively new intelligence committees as the key sources of congressional oversight in this area. Finally, there was the Intelligence Authorization Act, Fiscal Year 1991, the main legislative response to the Iran-Contra Affair, which defined “covert action” and memorialized in statute many of the practices and understandings that had grown up around the oversight of covert action since 1974.

Although it admits some variety as a function of presidential preference, the “systematic, institutionalized process” underpinning covert action is designed to evaluate “effectiveness, risk, and policy adherence.” Former officials who have participated in the finding process have offered insights as to how it is structured and what core questions need to be answered before action can be authorized. For example, former Secretary of State Cyrus Vance remarked that “it should be the policy of the United States to engage in covert actions only when they are absolutely essential to the national security.” Meanwhile, former CIA Director William Webster established four separate criteria for determining whether a proposed covert action ought to be undertaken: (1) “Is it legal?”; (2) “Is it consistent with American foreign policy, and, if not, why not?”; (3) “Is it consistent with American values?”; and (4) “If it becomes public, will it make sense to the American people?”
How well the finding process works in practice remains an open question. Some contend that it embodies an inherent short-term bias. Others argue that the effectiveness of the finding turns on the degree of specificity or generality in the relevant documents. For example, the recently issued Senate report on CIA interrogation suggests that the CIA's claimed authority to carry out its detention and interrogation activities was rooted in a Memorandum of Notification that was signed on September 17, 2001. But the Memorandum did not mention detention or interrogation at all, let alone the specific techniques that occasioned the report. It thus remains an open question what President Bush knew, and when, about the various interrogation techniques. That said, recent reporting on the drone program suggests that the President devotes extensive personal and White House attention to its oversight, approving targeting lists and weighing in on individual cases on the advice of senior policy and legal advisors. Flawed and contested though the process has been, covert-action findings have fortified the degree to which certain intelligence activities are exposed to oversight and have promoted serious consideration of the potential costs and benefits of pursuing inevitably controversial courses of action.

It makes sense to transpose the regulatory regime that has grown up around covert action to the world of intelligence collection. A first step in this process appears to have been taken already. The public portion of PPD-28 states that its classified annex mandates that “determinations about whether and how to conduct signals intelligence activities must carefully evaluate the benefits to our national interests and the risks posed by those activities.” Intimations of the finding process can be picked up here. The PRG's Recommendation 16 is that “the President should create a new process requiring high-level approval of all sensitive intelligence requirements and the methods the Intelligence Community will use to meet them.” Furthermore, as noted above, a select number of highly sensitive collection efforts already receive high-level White House attention. But a more formal institutionalization culminating in a legal requirement would allow for the realization of the full benefits of this sort of oversight technology. At a minimum, a finding ought to establish that “for any individual line of intelligence-gathering ... there is no reasonable alternative way of acquiring the information from less sensitive or non-secret sources, thereby avoiding all the possible moral hazards and trade-offs that collecting secret intelligence may involve.”

The most straightforward (and politically frictionless) way to accomplish this goal would be to amend Executive Order 12,333. Meanwhile, in order to trigger congressional oversight under existing legislation, it would be sufficient to deem any collection program that met the criteria for presidential review to be a “significant anticipated intelligence activity.” To be sure, structuring a process to handle intelligence collection, as opposed to covert action, requires careful thought. At the outset, there is a question of what the proper “unit of analysis” ought to be when it comes to surveillance. Should specific intelligence techniques or platforms warrant presidential attention? Or is surveillance of especially sensitive targets a more appropriate focus? Would all surveillance of foreign leaders, for example, require presidential sign-off? Or only the surveillance of the leaders of close allies? Making matters more complicated is the fact that unlike covert action, which is principally undertaken by the CIA, surveillance is a core function of all intelligence agencies and is vastly resourced. Only a small fraction of the CIA recruitment approaches, NSA collection orders, or National Reconnaissance Office imagery shots could be handled through a finding process without overwhelming the White House and undercutting a core motivation for a finding in the first place: meaningful engagement by the President and his senior staff with the potential costs and benefits of the underlying activity. One way to begin to get traction on the problem and to define a manageable universe of items that would receive attention in the White House would be to establish a set of presumptions that certain kinds of collection (say, on certain allies and foreign leaders, or employing certain means) would be potentially elevated for presidential attention.
Implementing a finding process might necessitate some changes in White House staffing in order to ensure the availability of the relevant capacity to assess the risks and rewards of contemplated surveillance practices. In particular, such a process may entail augmenting the NSC’s intelligence component in order to compensate for the reality that “agency officials frequently possess subject-specific skills and knowledge that the White House lacks.” Currently, the NSC’s intelligence directorate is small. By enlarging it, presidential intelligence can gain a necessary foothold within the nerve center of foreign policy and national security decisionmaking. But presidential intelligence is not a matter of hiring more spies in downtown Washington to oversee the work of spies in suburban Virginia and Maryland. Rather, it is imperative to expose intelligence programs to the generalist review that is characteristic of the NSC’s interagency process. It may well be that some amount of collection on allies is, in fact, vital to American security or that eroding the bottom lines of technology firms is a cost worth absorbing in some cases. But the critical task of clarifying the line where the advantages of collection are outweighed by the potential political and economic fallout in the aftermath of exposure would be greatly enhanced by a finding process that culminated in presidential sign-off. For the majority of intelligence collection efforts that will not qualify for the presidential finding process adumbrated above, the Office of the Director of National Intelligence could perform some amount of centralized review for overall coherence. It is arguably already accomplishing some of this work through the so-called “Mission Managers.” But inevitably, a great deal of discretion will remain in the hands of the various intelligence agencies.

B. Appointing Intelligence Officials

Moe and Wilson write that “[p]residents politicize by using their appointment authority to place loyal, ideologically compatible people in pivotal positions in the bureaus, the departments, and, of course, the OMB and other presidential agencies whose job it is to exercise control.” But politicization (in this sense) is a far more complicated strategy for the President to execute in the intelligence domain. Whereas the NSC supplies the institutional foundation for a centralizing tendency in intelligence oversight, the intelligence bureaucracy has been notably lacking in political appointments. To be sure, Presidents have placed political allies in delicate intelligence posts. One need think no further than President Reagan appointing his campaign manager Bill Casey, or President Obama tapping his national security confidant John Brennan, to run the CIA. The current deputy director of the CIA and his immediate predecessor are also administration veterans, as opposed to intelligence insiders. Nor has President Obama shied away from making changes in senior intelligence posts, as when he accepted the resignations of General David Petraeus at the CIA and Admiral Blair at the ODNI. But the trend does not generalize past these high-profile appointments and departures. By and large, intelligence agencies, notwithstanding the fact that they are quintessential executive agencies on the traditional administrative law metric, have successfully resisted politicization.

Currently, there are fewer than twenty Senate-confirmed presidential appointees across the sprawling intelligence state. For example, in the CIA, there are three such posts-- the Director, Inspector General, and General Counsel--only one of which has managerial responsibility for the organization. Within the FBI (which plays a key role as a domestic intelligence service in addition to its role in federal law enforcement), there is only the Director, who also enjoys a statutory ten-year term. While the President may fire the FBI Director in the middle of his term and appoint a replacement (and occasionally has ), there are political costs associated with doing so. President-elect Obama went so far as to pledge to “insulate the Director of National Intelligence from political pressure by giving the DNI a fixed term, like the Chairman of the Federal Reserve.” But there is no evidence that he has pursued that strategy since entering office.
To be sure, there are more officials who are appointed by the President (without a Senate role), but not that many. Most are chosen by department heads with the concurrence of the DNI. Even the NSA Director has faced Senate confirmation in the past-- by the Armed Services Committee, not the Intelligence Committee-- only because he concurrently serves as a senior military officer and, of late, also as the four-star officer in charge of Cyber Command. Under new legislation, the Director of the NSA will face Senate confirmation and hearings before the Intelligence Committee. Still, in its lack of politicization, the intelligence community remains an outlier in Washington. For purposes of comparison, within the executive branch there are anywhere from 1200 to 1400 posts that require presidential nomination and Senate confirmation. At the EPA alone there are a dozen such positions, and in the Pentagon there are approximately fifty.

Heightened politicization would accentuate the President's ability both to extract information from the intelligence bureaucracy and to ensure that his oversight preferences (or, for that matter, those of the FISC) are, in fact, implemented. The proliferation of lawyers in the various spy agencies provides a good template for thinking about the value of an increase in the number of political appointees. Just as the growing number of lawyers has promoted a culture of fidelity to law, an increase in the number of presidential appointees would foster greater attentiveness to the President's policies. Furthermore, the pool of potential political appointees with relevant backgrounds (who are not themselves current or former intelligence officials) might be small. But as the nature of intelligence evolves, individuals with training in a host of disciplines and professions ranging from area studies to data science would seem to be good candidates for intelligence posts. That some (especially intelligence veterans) will resist this move on the grounds that it will politicize intelligence should not be decisive. As discussed above, presidential control will always run some risk of eroding expert judgment. This concern is heightened in the context of the politicization of posts that demand neutral decisionmaking. But the sorts of managerial and leadership positions in intelligence that I advocate politicizing inevitably comingel normative judgments with intelligence assessments.

The paucity of Senate-confirmed posts is also a missed opportunity for congressional oversight. Here, too, caution must be exercised, because under conditions of hyperpartisanship, the requirement of Senate confirmation might be a recipe for deadlock on appointments and an invitation to grow the ranks of “acting” officials. But it remains the case that with fewer posts that require Senate confirmation come fewer opportunities for the public to learn about the career trajectories and backgrounds of senior intelligence personnel. Confirmation hearings of this sort would function as instances of meta-transparency where the public would be able to assess not the nature of intelligence collection as such but the nature of the people who practice it.

CONCLUSION

Almost fifty years ago, Professor Aaron Wildavsky offered that “[t]he United States has one President, but it has two presidencies; one presidency is for domestic affairs, and the other is concerned with defense and foreign policy.” For some time that claim has been off-target with respect to large swaths of the national security state, which have been on a convergence course with the ordinary regulatory state. But, until very recently, Wildavsky's observation retained some of its descriptive accuracy with respect to the intelligence community--specifically, as to the ways that spy agencies gather intelligence. Even as the President came to loom large in just about every other major area of policymaking, presidential involvement in the domain of intelligence collection remained episodic and minimal. That, too, is now changing. While the CIA, NSA, FBI, and every other spy agency each carries out a particular mission and maintains
a distinctive organizational culture, the intelligence community collectively is more than ever of a piece with the balance of government, in terms of the political and economic forces that affect it and the oversight methodologies and institutions that constrain it.

With intelligence having rejoined the regulatory mainstream after an extended hiatus, there is a lot of catching up that needs to be done. Building on previous work, this Article has emphasized the ways in which concepts and scholarly insights that have come to define administrative law are ripe for export to the intelligence bureaucracy. But in the future, interdisciplinary insights ought to flow in the other direction as well. The administrative state increasingly has things to learn from the intelligence world. For example, the intelligence community has had extensive experience with the legal and policy issues implicated by the revolution in “Big Data.” As the delivery of health care and education, for example, come to depend upon Big Data, regulatory agencies could learn a great deal from spy agencies. Furthermore, in an age of mounting threats in cyberspace, regulators from the “ordinary” administrative state--not to mention private actors--can and must learn from, and interact with, intelligence and national security agencies. But setting these emerging issues aside for future research, the key insight for now is that a meaningful and durable dialogue between the intelligence and regulatory states has begun.

In this Article, I have described a set of processes by which the intelligence community has been presidentialized, and have expressed qualified optimism that the trend will promote more effective, accountable, and rights-protective intelligence collection practices. Under conditions of robustly implemented presidential intelligence, the indiscriminate collection of American metadata premised on a secret, dubious statutory interpretation and the gratuitous eavesdropping on friendly foreign leaders' cellphone conversations will be less likely to come to pass again. At a minimum, the White House will have to devote serious attention to the potential upsides of these otherwise costly efforts before choosing to embark on them.

Footnotes

a1 Professor of Law, Faculty Director, Center on Law and Security, New York University School of Law. I would like to thank William Banks, Rachel Barkow, Gabriella Blum, Philip Bobbitt, Robert Chesney, Noah Feldman, John Ferejohn, Zachary Goldman, Jack Goldsmith, Ryan Goodman, Philip Heymann, Gavin Hood, Aziz Huq, Samuel Issacharoff, Michael Leiter, Daryl Levinson, Richard Morgan, Trevor Morrison, Matthew Olsen, Richard Pildes, David Pozen, Richard Revesz, and Stephen Slick, as well as workshop participants at New York University School of Law, Columbia Law School, and Hofstra Law School for helpful comments and suggestions. David Hoffman, Tyler Infinger, Nishi Kumar, Stephanie Spies, and Timothy Sprague provided excellent research assistance. Gretchen Feltes furnished characteristically superb library assistance.


2 OIRA is a component of the Office of Management and Budget (OMB), which functions as a clearinghouse for major rules. *See generally* RICHARD L. REVESZ & MICHAEL A. LIVERMORE, RETAKING RATIONALITY: HOW COST-BENEFIT ANALYSIS CAN BETTER PROTECT THE ENVIRONMENT AND OUR HEALTH (2008).


5 *See Aziz Huq, Imperial March, DEMOCRACY* (Winter 2008), http://www.democracyjournal.org/7/6571.php [http://perma.cc/6GHR-ENEC] (“From one angle, the Bush Administration's freewheeling unilateralism when it comes to interrogation and detention is merely the dark side of Clinton's exuberant, and often celebrated, unilateral use of executive agencies.”). A similar claim could be sustained with respect to President Obama's significant involvement in drone strikes. *See*

6 See Member Agencies, INTELLIGENCE.GOV, https://www.intelligencecareers.gov/icmembers.html [https://perma.cc/2FRZ-E9W7] (listing the seventeen separate organizations that form the “Intelligence Community”).

7 S. Res. 21, 94th Cong. (1975) (establishing a “select committee of the Senate to conduct an investigation and study with respect to intelligence activities carried out by or on behalf of the Federal Government,” later called the “Church Committee” after its chairman, Senator Frank Church). For the report issued by the Church Committee, see FINAL REPORT OF THE SELECT COMMITTEE TO STUDY GOVERNMENTAL OPERATIONS WITH RESPECT TO INTELLIGENCE ACTIVITIES, S. REP. NO. 94-755 (1976).

8 H.R. Res. 138, 94th Cong. (1975), replaced and expanded by H.R. Res. 591, 94th Cong. (1975) (establishing a committee that came to be known as the “Pike Committee” after its chairman, Representative Otis Pike, to parallel the Senate’s “Church Committee”). For the report issued by the Pike Committee, see RECOMMENDATIONS OF THE FINAL REPORT OF THE HOUSE SELECT COMMITTEE ON INTELLIGENCE, H.R. REP. NO. 94-833 (1976).


12 President Ford created the Intelligence Oversight Board with Exec. Order No. 11,905, 3 C.F.R. 90 (1977). It was later superseded by President Carter’s Exec. Order No. 12,036, 3 C.F.R. 112 (1979). President Clinton made the Intelligence Oversight Board a part of what was then called the President’s Foreign Intelligence Advisory Board (PFIAB) with Exec. Order No. 12,863, 3 C.F.R. 632 (1994).


15 There are some exceptions. In a short essay in the Harvard Journal on Legislation, James Baker (who has served as a senior intelligence lawyer in government and is currently General Counsel of the Federal Bureau of Investigation (FBI)) expressed the view that “it is first and foremost the President’s responsibility to conduct oversight of intelligence activities.” James A. Baker, Intelligence Oversight, 45 HARV. J. ON LEGIS. 199, 203 (2008).

16 See K.G. Robertson, The Politics of Secret Intelligence--British and American Attitudes, in BRITISH AND AMERICAN APPROACHES TO INTELLIGENCE 244, 249 (K.G. Robertson ed., 1987) (noting that, compared with overseas counterparts, American spy agencies enjoy “greater independence ... from direct political control”).


18 See Loch K. Johnson, The CIA and the Question of Accountability, in ETERNAL VIGILANCE?: 50 YEARS OF THE CIA 178, 180 (Rhodri Jeffreys-Jones & Christopher Andrew eds., 1997) (“Nor did the Executive Office of the Presidency (EOP) offer reliable accountability over the intelligence establishment that sprawled beneath the White House in the organizational
charts of the federal government.”). For a diagnosis of the power wielded by the national security bureaucracy, see generally Glennon, supra note 17.

19 My claims are not limited to any particular intelligence agency or collection platform. Whereas many of the post-Snowden developments that I document are particularly focused on electronic surveillance, the conceptual issues they implicate generalize to other intelligence disciplines, such as human intelligence. Furthermore, though critical attention regarding electronic surveillance has centered on the role of the National Security Agency (NSA), other organizations like the FBI have been deeply involved. See, e.g., Charlie Savage, F.B.I. Is Broadening Surveillance Role, Report Shows, N.Y. TIMES (Jan. 11, 2015), http://www.nytimes.com/2015/01/12/us/politics/beyond-nsa-fbi-is-assuming-a-larger-surveillance-role-report-shows.html (referring to a recently declassified--though still redacted--2008 report by the Department of Justice Inspector General assessing the role of the FBI in surveillance under the FISA Amendments Act). In particular, the FBI has been involved in the administration of the telephony metadata program under section 215 of the Patriot Act. See Michael Isikoff, FBI Sharply Increases Use of Patriot Act Provision to Collect US Citizens' Records, NBC NEWS (June 11, 2013, 11:42 AM), http://investigations.nbcnews.com/_news/2013/06/11/18887491-fbi-sharply-increases-use-of-patriot-act-provision-to-collect-us-citizens-records [http://perma.cc/L6N6-UHA7].

20 In thinking about the ways that intelligence has previously defied the norms of the administrative state, and in contemplating the path by which that exceptionality is now under pressure, I am indebted to the scholarship of Professor Rachel Barkow, who has questioned the nonapplicability of administrative law norms to the world of criminal law. See, e.g., Rachel E. Barkow, Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law, 61 STAN. L. REV. 869 (2009).

21 The emergence of presidential intelligence may be better thought of as a reemergence in that previous administrations attempted, but ultimately failed, to get the project off the ground. It is suggestive that Executive Order 12,333, 3 C.F.R. 200 (1982) (governing the intelligence community), never served as a font of centralized control on par with Executive Order 12,291, 3 C.F.R. 127 (1982) (requiring agencies to employ cost-benefit analysis), perhaps because the Iran-Contra scandal impaired the Reagan White House's ability to centralize control of intelligence. See GRIFFIN B. BELL WITH RONALD J. OSTROW, TAKING CARE OF THE LAW 139-41 (1982) (noting that in “the first months of the Reagan Administration ... [t]he Heritage Foundation ... proposed undoing virtually all intelligence reform measures,” including “doing away with the” FISC, id. at 139, but concluding that the Reagan Administration was ultimately unable to realize these ambitions).

22 Cf. ROY GODSON, DIRTY TRICKS OR TRUMP CARDS: U.S. COVERT ACTION AND COUNTERINTELLIGENCE 246 (1995) (“In terms of separation of powers, the world of U.S. intelligence has been ‘normalized.’”); Gregory F. Treverton, Intelligence: Welcome to the American Government, in INTELLIGENCE: THE SECRET WORLD OF SPIES, AN ANTHOLOGY 347 (Loch K. Johnson & James J. Wirtz eds., 3d ed. 2011) (noting that, as judged by the way in which congressional oversight of intelligence functions, the intelligence community belongs to the mainstream of American government).


26 See infra pp. 671-72.

The Office of Science and Technology Policy advises the President “on the effects of science and technology on domestic and international affairs.” See Office of Science and Technology Policy: About OSTP, WHITE HOUSE, https://www.whitehouse.gov/administration/eop/ostp/about [https://perma.cc/H5JZ-3S89]. Notably, one of its divisions is focused on national security and international affairs. See id.


See, e.g., Kent Roach, Review and Oversight of National Security Activities and Some Reflections on Canada's Arar Inquiry, 29 CARDOZO L. REV. 53, 55 (2007) (“[T]here may be much to be said for separating the processes of oversight and review ... [for the] efficacy of national security activities and ... [for their] propriety.”); cf. Baker, supra note 15, at 200-01 (“When it comes to conducting oversight of the United States intelligence community ... it seems that our goals should include ensuring that taxpayers' funds are spent appropriately and efficiently on programs and activities that produce useable intelligence information; that intelligence activities are effective in protecting the United States and its interests from foreign threats; and that intelligence activities are conducted in a lawful manner at all times.”).

It is worth heeding Professor Amy Zegart's caution that the very meaning of the word “oversight” is contested, with some arguing that it refers to a process of holding hearings and demanding accounts, others emphasizing substantive results in agency behavior and outputs, and still others reading into the term a normative requirement that the agency perform effectively in the national interest. See Amy B. Zegart, Agency Design and Evolution, in THE OXFORD HANDBOOK OF AMERICAN BUREAUCRACY 207 (Robert F. Durant ed., 2010).


Arguing for presidential intelligence does not entail dismissing or downplaying the ongoing significance of traditional oversight institutions or arguing for their demise. Judicial review, congressional oversight, and internal-executive branch checks (emanating from offices of general counsel, compliance chiefs, and various civil liberties-focused bodies) all have important, ongoing roles to play in the complex undertaking of intelligence oversight. Inasmuch as presidential intelligence tends to “crowd out” other overseers—a claim that I tackle head on—that ought to count against the advisability of the project. But this dynamic is certainly not inevitable. In fact, properly conceived and designed, presidential intelligence ought to contribute to the effectiveness of controls exercised by other overseers.

See Thomas O. Sargentich, The Emphasis on the Presidency in U.S. Public Law: An Essay Critiquing Presidential Administration, 59 ADMIN. L. REV. 1, 35-36 (2007) (criticizing the tendency of the presidential administration literature to mythologize the capacity of the White House, noting that its “accountability and effectiveness claims present a picture of the President as a white knight uniquely able to vindicate the public interest”).

Cf. GOLDSMITH, supra note 14, at 210 (“To say that the presidential [accountability system] helped generate a consensus about the counterterrorism policies the President can legitimately use does not, unfortunately, mean that it generated the right policies—the ones best designed to prevent terrorist attacks while ... preserving other values as much as possible.”). It has been repeatedly observed that pendulum swings in intelligence matters are pronounced. To take a striking example, on 9/11, former Secretary of State James Baker gave a television interview blaming the intelligence failure on the Church Committee, a vast overstatement that nevertheless conveyed a deep truth about perceptions in national security politics. ABC Sept. 11, 2001
I take up the matter of oversight of covert action later in the Article in order to draw a sustained analogy between that process and the oversight of intelligence gathering that figures prominently in this Article. See infra section IV.A, pp. 706-12.


Indeed, the very first published leak was of a FISC order compelling Verizon to turn over extensive metadata to the government. See Verizon Forced to Hand Over Telephone Data--Full Court Ruling, THE GUARDIAN (June 5, 2013, 7:04 PM), http://www.theguardian.com/world/interactive/2013/jun/06/verizon-telephone-data-court-order [http://perma.cc/QV5Z-8ANC].


That committee was the President's Review Group on Intelligence and Communications Technologies. See About the Review Group on Intelligence and Communications Technologies, OFFICE DIRECTOR NAT'L INTELLIGENCE, http://www.dni.gov/index.php/intelligence-community/review-group [http://perma.cc/L3K3-DX9U].


The line between the two phenomena is not always sharply drawn. Consider the recent career of John Brennan, for example. He served as President Obama's counterterrorism advisor, deepening a close personal bond with the President that he first forged as a campaign advisor. More recently, as CIA Director, Brennan has maintained those close ties. See Peter Baker & Mark Mazzetti, Brennan Draws on Bond with Obama in Backing C.I.A., N.Y. TIMES (Dec. 14, 2014), http://www.nytimes.com/2014/12/15/us/politics/cia-chief-and-president-walk-fine-line-.html (“[I]n the 67 years since the C.I.A. was founded, few presidents have had as close a bond with their intelligence chiefs as Mr. Obama has forged with Mr. Brennan.”).

The analogy does not work on every level. For example, the sense in which presidential administration represents, in part, an effort to energize agencies suffering from regulatory lethargy has no obvious corollary in the intelligence domain. For this reason, presidential intelligence may (more so than presidential administration) tend to skew “antiregulatory,” which is to say, antisurveillance. That said, once the presidential intelligence “game” is played repeatedly over the coming years, agencies may become more timid, creating a new equilibrium and necessitating a more catalytic approach from the White House.

Politicization in the pejorative sense that intelligence officials use the term, see infra section III.B.1, pp. 692-97, bears no resemblance to the way in which Moe deploys the term to refer to a process of presidential control that plays out within the agencies themselves, rather than within the White House. See infra notes 93-97 and accompanying text.


In practice, many of these intelligence functions are inevitably, and perhaps also increasingly, comngled. See infra p. 695. But standard intelligence texts continue to maintain these distinctions, and the intelligence profession continues to adhere to them.


During the 1990s, in the aftermath of the fall of the Berlin Wall and before 9/11, the relationship between the White House and the CIA became more tenuous. When a Cessna 150 landed on the White House lawn during the Clinton administration, a popular inside-the-Beltway joke suggested that it was piloted by CIA Director R. James Woolsey employing a desperate measure to get on the President's schedule. See AMY B. ZEGART, SPYING BLIND 71 (2007). This more remote relationship was profoundly altered by the events of 9/11.

See infra section IV.A, pp. 706-12. As a former CIA deputy director put it “Covert action is owned by the NSC and implemented by the CIA. Covert action is a tool of the President.” GENEVIEVE LESTER, WHEN SHOULD STATE SECRETS STAY SECRET 112 (2015) (quoting Stephen R. Kappes).


See MARSHALL CURTIS ERWIN, CONG. RESEARCH SERV., RL33715, COVERT ACTION: LEGISLATIVE BACKGROUND AND POSSIBLE POLICY QUESTIONS 1 (2013). For a discussion of subsequent legislative developments, see infra pp. 707-08.

See S. SELECT COMM. ON INTELLIGENCE, 113TH CONG., COMMITTEE STUDY OF THE CENTRAL INTELLIGENCE AGENCY’S DETENTION AND INTERROGATION PROGRAM 2-8 (Comm. Print 2014) [hereinafter DETENTION AND INTERROGATION REPORT].

Jo Becker & Scott Shane, Secret “Kill List” Proves a Test of Obama's Principles and Will, N.Y. TIMES (May 29, 2012), http://www.nytimes.com/2012/05/29/world/obamas-leadership-in-war-on-al-qaeda.html (“Mr. Obama has placed himself at the helm of a top secret 'nominations' process to designate terrorists for kill or capture ....


The FISDT of the National Security Advisor, BROOKINGS, http://www.brookings.edu/about/projects/archive/nsd/19991025 [http://perma.cc/W95J-BJQJ]. Other NSC posts, like the Assistant to the President for Homeland Security and Counterterrorism, also have extensive points of tangency with intelligence.

Historian of the NSC David Rothkopf has argued that the National Security Advisor position might have been even more influential in intelligence matters than it has proved to be, which would have obviated the need for a Director of National Intelligence. See DAVID J. ROTHKOPF, RUNNING THE WORLD: THE INSIDE STORY OF THE NATIONAL


See About the Review Group on Intelligence and Communications Technologies, supra note 45.


Steven C. Boraz, Executive Privilege: Intelligence Oversight in the United States, in REFORMING INTELLIGENCE 27, 32 (Thomas C. Bruneau & Steven C. Boraz eds., 2007).


The President’s Review Group couches one of its recommendations in terms of engaging senior policymaker input on issues that go beyond the first two tiers of NIPF priority. PRG, supra note 65, at 168 (“We recommend that: ... senior policymakers should review not only the requirements in Tier One and Tier Two of the National Intelligence Priorities Framework, but also any other requirements that they define as sensitive ....”).

And as the old adage has it, if everything is a priority, then nothing is.


ERWIN & BELASCO, supra note 82, at 10-11.

See MARK M. LOWENTHAL, INTELLIGENCE: FROM SECRETS TO POLICY 87 (6th ed. 2015).
The President and his senior staff are involved in signing off on highly sensitive technical collection decisions. See, e.g., BOB WOODWARD, VEIL: THE SECRET WARS OF THE CIA, 1981-1987, at 30 (2005) (tapping undersea cables required presidential sign off); see also Ryan Lizza, State of Deception: Why Won't the President Rein in the Intelligence Community?, NEW YORKER (Dec. 16, 2013), http://www.newyorker.com/magazine/2013/12/16/state-of-deception [http://perma.cc/2YKP-EL6K] (describing a briefing President Obama received in early February 2009 setting out substantial NSA compliance issues with FISC orders governing its metadata program, and the President's decision to proceed with the program when Judge Walton on the FISC was threatening to shut it down); Presidential Remarks on Signals Intelligence, supra note 42 (“I maintained a healthy skepticism toward our surveillance programs after I became President. I ordered that our programs be reviewed by my national security team and our lawyers, and in some cases I ordered changes in how we did business.”).

HARRY HOWE RANSOM, THE INTELLIGENCE ESTABLISHMENT 247 (1970); see also Kenneth deGraffenreid, Intelligence and the Oval Office, in 7 INTELLIGENCE REQUIREMENTS FOR THE 1980'S: INTELLIGENCE AND POLICY 9, 16 (Roy Godson ed., 1986) (“If a president is interested in having a closer look at one issue than at another, he ought to, even if it means a reordering or restructuring of the intelligence community's collection and analytic efforts.”).


Id. at 19 (“The continuing problem for presidents, though, is that they have too little control, not too much, and they need to build an institution that helps them do a better job of overcoming the tremendous obstacles to leadership the system places in their way. This is what the presidential team, the various presidential organizations, and the strategies of politicization and centralization are all about, and it is what the institutional presidency as a whole is all about. This is how presidents fight back: with structures that enhance their power.”).

Moe, supra note 46, at 157.

See Moe & Wilson, supra note 91, at 18-19.

Former OIRA Administrator Sally Katzen has remarked on how important it is to OIRA that it sits within the institutional presidency at OMB. See Sally Katzen, Correspondence, A Reality Check on an Empirical Study: Comments on “Inside the Administrative State,” 105 MICH. L. REV. 1497, 1498 (2007) (“[T]he centralized review of rule-makings is only one piece of presidential control. [OMB] presides over the whole mosaic, where review of rule-making occurs along side review of legislative proposals, review of Executive Orders, and, very importantly, review of resource (budget) decisions.”).

See Kagan, supra note 1, at 2310 (“More than any other player in the political system, the President is in practice, even if not in constitutional theory, responsible for governance.”).

See Moe & Wilson, supra note 91, at 18.


ZEGART, supra note 32, at 115-16. Zegart's focus is on congressional oversight, but her observation generalizes beyond Capitol Hill.

Cf. Johnson, supra note 18, at 194 (discussing the lobbying efforts undertaken by the CIA on its own behalf).
Not that there are no bureaucratic incentives pointing in this direction. For example, former CIA acting General Counsel John Rizzo has written that he should have sought greater congressional oversight of the legal foundations of the CIA's interrogation program in order to more effectively distribute political risk. See JOHN RIZZO, COMPANY MAN: THIRTY YEARS OF CONTROVERSY AND CRISIS IN THE CIA 200-01 (2014).


Michael Hayden, as NSA Director, embraced the PSP. He was subsequently tapped by President Bush to serve as Deputy Director of National Intelligence and then Director of the CIA. See Dan Eggen & Walter Pincus, Campaign to Justify Spying Intensifies, WASH. POST (Jan. 24, 2006), http://www.washingtonpost.com/wp-dyn/content/article/2006/01/23/AR2006012300754_pf.html [http://perma.cc/NBM6-UKV6]. Then-DNI John Negroponte did not offer a public defense of the program, leaving Hayden in charge. See RICHARD A. POSNER, COUNTERING TERRORISM: BLURRED FOCUS, HALTING STEPS 71-72 (2007).


For a thoughtful insider's view of this process, see MICHAEL ALLEN, BLINKING RED: CRISIS AND COMPROMISE IN AMERICAN INTELLIGENCE AFTER 9/11 (2013).

See RICHARD A. POSNER, PREVENTING SURPRISE ATTACKS: INTELLIGENCE REFORM IN THE WAKE OF 9/11, at 140-42 (2005). This squares with Barron's assessment of OIRA. See Barron, supra note 98, at 1112 (“Far from imposing a presidential/political view of the world on top of an administrative/expert one, OIRA review is better conceptualized as instituting a new layer of technical (even neutral, bureaucratic) review, but one that is much more deregulatory in orientation because of the substantive inquiry that it requires OIRA analysts to undertake.”).

ALLEN, supra note 106, at 65. As 9/11 Commission Executive Director Philip Zelikow testified, “We recommended [locating the ODNI within] the executive office of the president because of the need for proximity to the president and the National Security Council.” Id. Ultimately the proposal was rejected out of concern that placement in the White House would politicize intelligence. See id.

So much so that an experienced Washington hand like Robert Gates writes that he declined President Bush's overture to become the first DNI when he read the statute and realized how weak the position was. In the intelligence lore of contemporary Washington, this moment sealed the fate of the ODNI. See Robert M. Gates, Opinion, Racing to Ruin the CIA, N.Y. TIMES (June 8, 2004), http://www.nytimes.com/2004/06/08/opinion/racing-to-ruin-the-cia.html (suggesting of the as-yet-to-be-created DNI that the “intelligence czar would, in fact, be an intelligence eunuch”); David Ignatius, Opinion, Gates's Next Mission, WASH. POST (Aug. 7, 2008), http://www.washingtonpost.com/wp-dyn/content/article/2008/08/06/AR2008080602511.html [http://perma.cc/M63F-QK8T].


Bobby Ghosh, Overseas Turf War Between the CIA and DNI Won't Die, TIME (Nov. 6, 2009), http://content.time.com/time/nation/article/0,8599,1936129,00.html.
112 See H. COMM. ON OVERSIGHT AND GOVT REFORM, 112TH CONG., POLICY AND SUPPORTING POSITIONS 177 (Comm. Print 2012) [hereinafter PLUM BOOK].

113 LOCH K. JOHNSON, NATIONAL SECURITY INTELLIGENCE 178 (2012).

114 S. SELECT COMM. TO STUDY GOVERNMENTAL OPERATIONS WITH RESPECT TO INTELLIGENCE ACTIVITIES, BOOK II: INTELLIGENCE ACTIVITIES AND THE RIGHTS OF AMERICANS, S. REP. NO. 94-775, at v (1976). Professor Margo Schlanger argues that there is a causal relationship between the dearth of policy-based review and the availability of legalist oversight in that the “relentless focus on rights and compliance and law ... has obscured the absence of what should be an additional focus on interests, or balancing, or policy.” See Margo Schlanger, Intelligence Legalism and the National Security Agency's Civil Liberties Gap, 6 HARV. NATL SECURITY J. 112, 118 (2015).

115 See RIZZO, supra note 101, at 44 (recounting a three-decades-long career as a lawyer at the CIA).

116 See Schlanger, supra note 114, at 190-91.

117 See Sinnar, supra note 32, at 1032 (“I identify five dimensions of rights oversight consistent with IGs' statutory mandate and analyze how IG reviews both contributed to these objectives and sometimes failed to do so: increasing transparency, identifying rights violations and wrongful conduct, providing relief for victims, holding government officials accountable for abuses, and revising agency rules to prevent future abuse.”).

118 See, e.g., FBI, DOMESTIC INVESTIGATIONS AND OPERATIONS GUIDE (2011).


122 The President's Intelligence Advisory Board and Intelligence Oversight Board, WHITE HOUSE, https://www.whitehouse.gov/administration/eop/piab [https://perma.cc/A82D-5TH4].

123 The PCLOB came to enjoy certain measures of independence after an early brouhaha that centered on top-down White House control. See Lanny Davis, Why I Resigned from the President's Privacy and Civil Liberties Oversight Board--And Where We Go from Here, THE HILL (May 18, 2007, 2:15 PM), http://thehill.com/blogs/pundits-blog/the-administration/34214-why-i-resigned-from-the-presidents-privacy-and-civil-liberties-oversight-board--and-where-we-go-from-here- [http://perma.cc/WPD2-ZW88]. It is still plagued by political problems. See Shirin Sinnar, Institutionalizing Rights in the National Security Executive, 50 HARV. C.R.-C.L. L. REV. 289, 316 (2015) (“[T]he ideological divides that contributed to the Board's long dormancy have resurfaced to thwart consensus on liberty-security questions.”). And its overall impact remains uncertain; Schlanger has recently referred to the board as “a blue-ribbon-commission type organization with no enforcement or other executive function.” Schlanger, supra note 114, at 166.

124 See, e.g., PRIVACY & CIVIL LIBERTIES OVERSIGHT BD., REPORT ON THE SURVEILLANCE PROGRAM OPERATED PURSUANT TO SECTION 702 OF THE FOREIGN INTELLIGENCE SURVEILLANCE ACT 104 (2014) (“Because the oversight mandate of the Board extends only to those measures taken to protect the nation from terrorism, our focus in this section is limited to the counterterrorism value of the Section 702 program, although the program serves a broader range of foreign intelligence purposes.”).

125 See Schlanger, supra note 114, at 172-88.


Notably the PCLOB has committed to studying surveillance conducted pursuant to 12,333. See Press Release, Privacy & Civil Liberties Oversight Bd., PCLOB Announces Its Short-Term Agenda (Sept. 3, 2014), http://www.pclob.gov/newsroom/20140807.html [http://perma.cc/RP6Q-ZP9R]. Informing the debate is an OLC memo (revealed by Senator Sheldon Whitehouse) that expresses the view that Presidents are not bound by the terms of executive orders (including Executive Order 12,333) in that departures from orders are plausibly deemed modifications. See 153 CONG. REC. 33,492-94 (2007) (statement of Sen. Whitehouse).


See SWIRE, *supra* note 23.

See PRG, *supra* note 65, at 170; see also Goldsmith, *supra* note 131 (“Secret intelligence actions-- especially the ones that would most likely engender outrage, surprise, debate, or legal controversy--are increasingly difficult to keep secret.”).

The ACLU’s Jameel Jaffer rightly observed “that Congress isn’t the only forum in which surveillance reform can be achieved,” emphasizing that “technology companies whose cooperation the government needs in order to conduct surveillance have already taken multiple steps to limit government surveillance.” Jameel Jaffer, *There Will Be Surveillance Reform, JUST SECURITY* (Nov. 20, 2014, 11:15 AM), http://justsecurity.org/17622/surveillance-reform [http://perma.cc/8TTL-KNRF].


My argument does not depend on the motivations driving firms and allies. It rests solely on their power and their ability to lean on the White House to achieve reforms. As a descriptive matter, the fact of interest group pressure on the President concerning intelligence practices bears out the accuracy of the observation that the White House is itself a site of interest group contestation. See generally Nicholas Bagley & Richard L. Revesz, Centralized Oversight of the Regulatory State, 106 COLUM. L. REV. 1260 (2006) (“As an initial matter, it would be naïve to assume that the President is immune to public choice pressures. He is not.” Id. at 1305.).

See Loch K. Johnson, Congressional Supervision of America's Secret Agencies: The Experience and Legacy of the Church Committee, in INTELLIGENCE: THE SECRET WORLD OF SPIES, AN ANTHOLOGY, supra note 22, at 393, 394. Of course, it is not correct that major technology and telecommunications firms have been strangers to national security policymaking or politics until very recently. To take a striking example, the telecommunications firms fought hard to get immunity from civil liability built into the structure of the FISA Amendments Act of 2008. See EDWARD C. LIU, CONG. RESEARCH SERV., RL34600, RETROACTIVE IMMUNITY PROVIDED BY THE FISA AMENDMENTS ACT OF 2008 (2008). And some aspects of the current political economy remind thoughtful observers of a prior generation's so-called crypto wars. See Joris V.J. van Hoboken & Ira S. Rubinstein, Privacy and Security in the Cloud: Some Realism about Technical Solutions to Transnational Surveillance in the Post-Snowden Era, 66 ME. L. REV. 487, 500-03 (2014) (describing a standoff between the technology industry and the national security state during the Clinton Administration over commercial uses of encryption technology).

The story of the emergence of presidential intelligence could itself be recast as a successful capture story, with the technology firms and foreign allies doing the capturing. See, e.g., Michael A. Livermore & Richard L. Revesz, Regulatory Review, Capture, and Agency Inaction, 101 GEO. L.J. 1337, 1340 (2013) (“Capture describes situations where organized interest groups successfully act to vindicate their goals through government policy at the expense of the public interest.”). That is certainly how many intelligence insiders who oppose the influence that technology firms currently wield see it.


For example, the structure of surveillance under Section 702 of FISA involves the issuance of directives by the government to private actors. 50 U.S.C. § 1881a(h) (2012) (“[T]he Attorney General and the Director of National Intelligence may direct, in writing, an electronic service provider ....”).


Sanger, supra note 41.


Musil, supra note 143.


See Microsoft Corp. v. United States, No. 14-2985 (2d Cir. filed Aug. 12, 2014).


60 Minutes: FBI Director on Privacy, Electronic Surveillance (CBS television broadcast Oct. 12, 2014), http://www.cbsnews.com/news/fbi-director-james-comey-on-privacy-and-surveillance [http://perma.cc/3SGJ-TGCK]. In the aftermath of the recent terrorist attacks in Paris, CIA Director John Brennan criticized the ways in which “hand-wringing” about government surveillance has led to legal and policy changes that, in turn, have made global counterterrorism efforts

159 Robert Hannigan, The Web is a Terrorist's Command-and-Control-Network of Choice, FIN. TIMES (Nov. 3, 2014, 6:03 PM), http://www.ft.com/intl/cms/s/2/c89b6c58-6342-11e4-8a63-00144feabdc0.html. Hannigan went on to call for “a new deal between democratic governments and the technology companies in the area of protecting our citizens.” Id.


161 Sometimes pressure from allies has merged with pressure from technology firms, as when Google Chairman Eric Schmidt spoke about his meeting with Chancellor Angela Merkel and her sense of outrage at surveillance practices that evoked her childhood experience in an East German surveillance state. See Nancy Scola, Google's Schmidt: Surveillance Fears Are “Going to End Up Breaking the Internet,” WASH. POST (Oct. 8, 2014), http://www.washingtonpost.com/blogs/the-switch/wp/2014/10/08/googles-schmidt-surveillance-fears-are-going-to-end-up-breaking-the-internet [http://perma.cc/GRH3-WU3A].


163 German Foreign Minister Summons US Ambassador over Merkel Spying Allegations, DEUTSCHE WELLE (Oct. 24, 2013), http://www.dw.de/german-foreign-minister-summons-us-ambassador-over-merkel-spying-allegations/a-17180294 [http://perma.cc/M7CK-5MP8]. Merkel commented on the allegations by saying that “trust needs to be re-established” with Washington. Id. German Defense Minister Thomas de Maiziere said it would be “really bad” if the allegations turned out to be true: “We can't simply return to business as usual … [but] the relations between our countries are stable and important for our future; they will remain that way.” Id.


Protests of NSA activity by foreign governments may be matters of politics, not principles. Recent reporting suggests that “the Merkel government knew of cooperation between the German foreign intelligence agency, the Bundesnachrichtendienst, and the American spy services, but withheld that information from a parliamentary committee assigned to investigate


Justin Huggler, *Germany to ‘Spy on US and UK Intelligence Gathering’ for the First Time in 45 Years*, THE TELEGRAPH (July 24, 2014, 4:57 PM), http://www.telegraph.co.uk/news/worldnews/europe/germany/10988939/Germany-to-spy-on-US-and-UK-intelligence-gathering-for-the-first-time-in-45-years.html. Interestingly, the implication of this newspaper article is that alleged U.S. spies in Germany were arrested without a concerted counterespionage program focused on American intelligence. See id.

181 Ken Dilanian, *AP Exclusive: CIA Halts Spying in Europe*, HUFFINGTON POST (Sept. 20, 2014, 12:09 AM), http://www.huffingtonpost.com/huff-wires/20140920/us-cia-europe-spying-pause [http://perma.cc/8QLF-NPQZ] (“Under the stand-down order, case officers in Europe largely have been forbidden from undertaking ‘unilateral operations’ such as meeting with sources they have recruited within allied governments. Such clandestine meetings are the bedrock of spying.”).
See, e.g., Conor Dougherty, Jay Edelson, the Class-Action Lawyer Who May Be Tech's Least Friended Man, N.Y. TIMES (Apr. 4, 2015), http://www.nytimes.com/2015/04/05/technology/unpopular-in-silicon-valley.html (noting how the Snowden leaks have improved the prospects of lawsuits seeking to hold technology firms accountable for privacy violations).

See Moe, supra note 46, at 144.

See supra pp. 649-50.

See supra pp. 649-50.


PPD-28, supra note 43. Commentators have reserved judgment about the ultimate impact of the PPD. One especially keen observer has written that it “represents an unprecedented change in U.S. intelligence policy, at least at the rhetorical level” even as “[t]he degree of substantive change that will follow from PPD-28 is less certain.” David S. Kris, On the Bulk Collection of Tangible Things, 7 J. NAT'L SECURITY L. & POL'Y 209, 289 (2014).

See Kagan, supra note 1, at 2290-99.


PPD-28, supra note 43. Among the significant policy changes ushered in by the directive and the accompanying speech, Presidential Remarks on Signals Intelligence, supra note 42, are (1) imposing a two-hop (rather than three-hop) standard on querying metadata; (2) recommending that the FISC, rather than the NSA, make findings about reasonable, articulable suspicion (it is unclear what authority the President employed to make this change, but the FISC seems to have assumed the responsibility notwithstanding the public letter by Chief Judge Bates suggesting the court was overburdened, Letter from Chief Judge John D. Bates, Director, Admin. Office of the U.S. Courts, to Patrick J. Leahy, Chairman, Comm. on the Judiciary (Aug. 5, 2014), http://online.wsj.com/public/resources/documents/Leahyletter.pdf [http://perma.cc/4FAE-YYAJ]); and (3) extending certain heightened privacy protections to foreign nationals. This last point makes sense as a direct response to the global pressures that have been brought to bear on the White House, both by allied governments and by technology firms with global customers. See supra section II.A, pp. 660-69.


PPD-28, supra note 43; see Laura K. Donohue, FISA Reform, 10 I/S: J.L. & POL'Y FOR INFO. SOCY 599, 612-13 (2014) (“PPD-28 ... lay[s] out the current principles guiding SIGINT, such as the integration of privacy and civil liberties
considerations in the collection of intelligence, limits on the collection of commercial information and trade secrets, and the tailoring of SIGINT to areas where the information is not otherwise available. The document restricts the use of bulk SIGINT data. It draws attention to ... minimization [procedures,] ... data security and access, data quality, and oversight.” (footnotes omitted)).

196 Presidential Remarks on Signals Intelligence, supra note 42.

197 Id.

198 PPD-28, supra note 43. Furthermore, PPD-28 calls for the creation of a post in the State Department to manage the diplomatic aspects of surveillance. Id.

199 Moe & Wilson, supra note 91, at 19 (“The president clearly has strong reasons for not wanting the State Department, the Defense Department, and other agencies to make their own foreign policy decisions.”).


201 Previously, President Obama had changed the name of the staff from the “National Security Council Staff” to the “National Security Staff,” or NSS. The purpose of the name change to NSS was to accommodate the merger of the Homeland Security Council and National Security Council staffs. See Michael D. Shear, Security Staff Getting Its Old Name Back, N.Y. TIMES (Feb. 10, 2014), http://www.nytimes.com/2014/02/11/us/politics/security-staff-getting-its-old-name-back.html.

202 This was not always so. See Johnson, supra note 18, at 181 (describing how, prior to the era of inquests into intelligence practices and greater accountability, “[m]any of the CIA's activities ... never received a thorough examination--or, in some cases, even approval--by the NSC”).

203 OIRA operates similarly. See Ryan Bubb, Comment, The OIRA Model for Institutionalizing CBA of Financial Regulation, 78 LAW & CONTEMP. PROBS., no. 3, 2015, at 47, 53 (“If the agency staff and [the OIRA staff] could not resolve an issue, it was elevated to the administrator of OIRA and the relevant senior political appointee at the promulgating agency to resolve. Failing their agreement, it went up the chain in the White House, and ultimately the president was indeed ‘the decider.’”).

204 Cf. PRG, supra note 65, at 169 (recommending the creation of an office lodged in the ODNI that reviews sensitive collection and that includes elements from nontraditional national security organizations “such as the National Economic Council, Treasury, Commerce, and the Trade Representative”); id. at 168 (recommending that senior policymakers from the federal agencies with responsibility for U.S. economic interests should participate in the review process because disclosures of classified information can have detrimental effects on those interests).

205 PPD-28, supra note 43.

206 See Moe & Wilson, supra note 91, at 19.

207 NEC Director Jeffrey Zients and OSTP Director John Holdren were coauthors of the White House's May 2014 report on big data. See PODESTA ET AL., supra note 75. Furthermore, the President, in his speech accompanying the issuance of PPD-28, noted his intention to appoint a “senior official at the White House to implement the new privacy safeguards that I have announced today.” Presidential Remarks on Signals Intelligence, supra note 42.


209 The pattern of striking a balance between employing the chief of staff and the national security advisor on matters that involve not only national security equities but also broader policy and political interests is longstanding. For example, Ken Duberstein, who served as Reagan's chief of staff toward the end of his presidency, had a fully worked-out modus operandi with then
National Security Advisor Colin Powell. The two of them would coordinate each morning as to which of the two might raise an issue with the President when it straddled the line between national security and larger policy and political concerns. See ROTHKOPF, supra note 70, at 256.


The View from the West Wing, ASPEN INST. (July 26, 2014), http://www.aspeninstitute.org/video/view-west-wing.

Id.


PRG, supra note 65, at 220.


See Barron, supra note 98, at 1111.

See, e.g., REVESZ & LIVERMORE, supra note 2.

See Stuart Shapiro, OIRA Inside and Out, ADMIN. L. REV., Special Edition 2011, at 135, 145 (noting that former OIRA Administrators “[S. Jay] Plager and [Sally] Katzen both cited the need for a cooperative relationship between agencies and OIRA, with Plager saying, ‘Agencies are not bad people’”).

Id. (“This repeated interaction is much less effective if the relationship between the desk officer and the agency is uniformly hostile.”).

Sidney A. Shapiro, Does OIRA Improve the Rulemaking Process? Cass Sunstein's Incomplete Defense, ADMIN. & REG. L. NEWS, Fall 2013, at 6, 6; see also Lisa Heinzerling, Inside EPA: A Former Insider's Reflections on the Relationship Between the Obama EPA and the Obama White House, 31 PACE ENVTL. L. REV. 325, 342 (2014) (“From my perspective, it was often hard to tell who exactly was in charge of making the ultimate decision on an important regulatory matter.”).

See, e.g., Nina A. Mendelson, Disclosing “Political” Oversight of Agency Decision Making, 108 MICH. L. REV. 1127, 1135 (2010) (“Presidential supervision clearly can make pragmatic contributions to agency decision making. A President can ensure that decision making among multiple federal agencies is coordinated. A President can provide direction and energy to agency officials. And centralized presidential supervision can counteract the tendency of an agency to take a ‘tunnel vision’ approach by bringing a broad perspective to agency prioritization and decision making.” (footnotes omitted)). Some scholars have questioned the capacity of OIRA to deliver on the goal of harmonization. See Lisa Schultz Bressman & Michael P. Vandenbergh, Inside the Administrative State: A Critical Look at the Practice of Presidential Control, 105 MICH. L. REV. 47, 75 (2006) (“As critics warn, OIRA review may not advance intra-agency coherence and interagency coordination at all or well enough.”).

**Id.** at 1855 (“A central goal of the OIRA process is to ensure that rulemaking agencies have access to the wide variety of perspectives that can be found throughout the executive branch.”). This is consistent with a scholarly emphasis on collaborative (as between agencies) regulation. See, e.g., William W. Buzbee, Recognizing the Regulatory Commons: A Theory of Regulatory Gaps, 89 IOWA L. REV. 1 (2003); Jody Freeman & Jim Rossi, Agency Coordination in Shared Regulatory Space, 125 HARV. L. REV. 1131 (2012).

Brad Plumer, Cass Sunstein on How Government Regulations Could Be a Lot Simpler, WASH. POST: WONKBLOG (June 12, 2013), http://www.washingtonpost.com/blogs/wonkblog/wp/2013/06/12/cass-sunstein-on-how-government-regulations-could-be-a-lot-simpler [http://perma.cc/CLY6-MUS7]. But see Bressman & Vandenbergh, *supra* note 223, at 50 (“OIRA review appears to advance inter-agency coordination somewhat, minimizing overlaps and conflicts between or among the regulations of different federal agencies. But OIRA review does not achieve what might be called ‘intra-agency coherence,’ which includes reducing redundancies, avoiding inconsistencies, and eliminating unintended consequences between or among the regulations of a particular agency. Thus, OIRA review fails to discharge one of the central purposes for which President Reagan created it and all subsequent presidents have maintained it.” (footnote omitted)).

**See** Amy Sinden, Formality and Informality in Cost-Benefit Analysis, 2015 UTAH L. REV. 93, 98-99. Some of the normative criticisms that have been lodged against CBA take aim at its more technical aspects. See, e.g., FRANK ACKERMAN & LISA HEINZERLING, PRICELESS: ON KNOWING THE PRICE OF EVERYTHING AND THE VALUE OF NOTHING (2004).

**See** POSNER, *supra* note 107, at 149 (advocating strong coordination of the intelligence agencies in respect to their collection function).


Presidential intelligence will naturally emphasize the costs and benefits to the United States and its citizens. But more cosmopolitan perspectives may also be required, for example, when the risks to allies are taken into account. Cf. Eric A. Posner & Cass R. Sunstein, Climate Change Justice, 96 GEO. L.J. 1565, 1572 (2008) (“We do not question the proposition that an international agreement to control greenhouse gases, with American participation, is justified, and all things considered, the United States should probably participate even if the domestic cost-benefit analysis does not clearly justify such participation.”)


For example, the then-Chairman of the Senate Foreign Relations Committee expressed frustration with Secretary of State John Kerry, who was testifying before the committee about the projected scale of military operations in Iraq and Syria but was not forthcoming about intelligence operations in the same area. As Senator Robert Menendez said to Secretary Kerry: “It is unfathomable to me to understand how this committee is going to get to those conclusions without understanding all of the elements of military engagement, both overtly and covertly.” Niels Lesniewski & Humberto Sanchez, *Before Approving ISIS War, Menendez Wants Intelligence Briefing*, ROLL CALL (Sept. 22, 2014, 3:50 PM), http://blogs.rollcall.com/wgdb/before-approving-isis-war-menendez-wants-intelligence-briefing [http://perma.cc/Z84M-KDYT].


Some question whether President Obama was ignorant of the tapping. See, e.g., *Embassy Espionage: The NSA's Secret Spy Hub in Berlin*, DER SPIEGEL (Oct. 27, 2013, 7:02 PM), http://www.spiegel.de/international/germany/cover-story-how-nsa-spied-on-merkel-cell-phone-from-berlin-embassy-a-930205-2.html [http://perma.cc/FY4H-C24S] (“Among the politically decisive questions is whether the spying was authorized from the top: from the US president.”). However, President Obama, backed by the NSA, has claimed that he would have halted the collection had he known about it. *NSA Says Obama Didn't Know Merkel's Phone Was Being Bugged*, AL JAZEERA AM. (Oct. 27, 2013, 10:52 PM), http://america.aljazeera.com/articles/2013/10/26/us-may-have-buggedangelamerkelsphonesince2002report. [http://perma.cc/S2FT-VMZ2].

The question of international law’s application to intelligence is complex and evolving. For a thoughtful framing of some of the issues surrounding disparate treatments of citizens and foreigners for purposes of bulk collection, see Ryan Goodman, *Should Foreign Nationals Get the Same Privacy Protections Under NSA Surveillance—or Less (or More)?*, JUST SECURITY (Oct. 29, 2014, 10:33 AM), http://justsecurity.org/16797/foreign-nationals-privacy-protections-nsa-surveillance-or-or-more [http://perma.cc/JF72-NVWE]. Professor Ashley Deeks has advanced the idea that some (limited, at least at first) international compact ought to govern this space. See Ashley Deeks, *An International Legal Framework for Surveillance*, 55 VA. J. INT'L L. 291 (2015); see also Kenneth Roth, *Obama & Counterterror: The Ignored Record*, N.Y. REV. BOOKS (Feb. 5, 2015), http://www.nybooks.com/articles/archives/2015/feb/05/obama-counterterror-ignored-record (“Brazil and Germany have led an initiative at the UN General Assembly to articulate global concern about the harm of mass surveillance for our right to privacy and other basic freedoms. In addition, there will be an effort at the UN Human Rights Council in March 2015 to create a special rapporteur on privacy who could elaborate global standards for communications regardless of where people are located.”).

As Zegart has argued, the CIA was founded with this sort of strategic outlook in mind. Over the years, it has drifted away from this core mission. See generally ZEGART, supra note 200.

See John C. Coates IV, *Cost-Benefit Analysis of Financial Regulation: Case Studies and Implications*, 124 YALE L.J. 882 (2015). But see Bubb, *supra* note 203, at 49 (arguing that cost-benefit analysis of financial regulation is “no more difficult--indeed, it might be less difficult-- than it is in many other domains in which it plays a central role”); Eric A. Posner & E. Glen Weyl, *Benefit-Cost Paradigms in Financial Regulation*, 43 J. LEGAL STUD. S1, S30 (2014) (recommending that “the president ... create a department within OIRA and give it the specific mission of coordinating [cost-benefit analysis] among the financial agencies”).

Amy B. Zegart, *The Domestic Politics of Irrational Intelligence Oversight*, 126 POL. SCI. Q. 1, 4 (2011). Zegart's focus is on congressional oversight, but some of her observations generalize more broadly across various oversight institutions.


See, e.g., COMM. ON RESPONDING TO SECTION 5(D) OF PRESIDENTIAL POLICY DIRECTIVE 28, NAT'L RESEARCH COUNCIL, REPORT ON BULK COLLECTION OF SIGNALS INTELLIGENCE: TECHNICAL OPTIONS (2015) (employing various “use cases” to assess the viability of software substitutes for bulk collection).


See Schlanger, supra note 114, at 113. One of the Board’s members filed a dissent in which she expressed the view that the Board ought to adopt a more policy-oriented, and less law-focused, outlook on metadata collection. See 215 PCLOB REPORT, supra note 244, at 210 (separate statement of Board Member Rachel Brand).

215 PCLOB REPORT, supra note 244, at 217 (separate statement of Board Member Elisebeth Collins Cook).

Joseph Fitsanakis, Secret Report Warns US Spy Mission Distorted by “War on Terror,” INTELNEWS.ORG (Mar. 22, 2013), http://intelnews.org/2013/03/22/01-1222 (reporting that a classified report compiled by President Obama’s Intelligence Advisory Board cautioned the President that the CIA and NSA had “been disabled by tunnel vision and operational fatigue in the pursuit” of al Qaeda and the focus on Islamic militancy).


This trend is noteworthy given that the intelligence establishment initially resisted the pivot to legalist oversight. See Margo Schlanger, A Cult of Rules: The Origins of Legalism in the Surveillance State, JUST SECURITY, (Nov. 5, 2014, 11:13 AM), http://justsecurity.org/17117/cult-rules-origins-intelligence-legalism [http://perma.cc/38KW-2GVU] (citing a drafter of Executive Order 12,333 for the proposition that there was “enormous pent-up hostility in the intelligence community toward lawyers and legalistic restrictions” and that the “attitude was not an invention of the Republican political appointees--who at that time were not yet that numerous--but permeated the career service”).


In a related vein, the ACLU’s Jameel Jaffer has criticized officials for touting (pre-Snowden, anyhow) the fact that the government's surveillance programs had been countenanced by all three branches of government. As Jaffer argued, “what it presented as a defense of the surveillance program was actually an indictment of our oversight system.” See Jameel Jaffer, Obama Is Cancelling the NSA Draget. So Why Did All Three Branches Sign Off?, THE GUARDIAN (Mar. 25, 2014, 10:09 AM), http://www.theguardian.com/commentisfree/2014/mar/25/obama-nsa-dragnet-phone-proposal-sign-off [http://perma.cc/YQF3-XCAH].


See DETENTION AND INTERROGATION REPORT, supra note 63.

Baker & Mazzetti, supra note 47 (quoting Professor David Cole).


See Jerry L. Mashaw, Accountability and Institutional Design: Some Thoughts on the Grammar of Governance, in PUBLIC ACCOUNTABILITY 115, 121 (Michael W. Dowdle ed., 2006) (setting out various types of public accountability and describing political accountability as a system in which “[t]op-level bureaucrats ... are responsible or accountable to an elected official ... for carrying out their discretionary functions in accordance with their political superiors' policies or ideological commitments”).

deGraffenreid, supra note 88, at 16.

See Kagan, supra note 1, at 2301 (“Some of this activity no doubt related more to strategies of public relations than of administrative governance. All methods of ‘going public,’ in the sense that political scientists use the term, aim to cultivate public support, and Clinton focused on this goal with equal or greater intensity than any of his predecessors. If, as I have indicated, his ‘going public’ strategy had a peculiarly administrative cast, a prime cause lay in his understanding that announcing new actions captured more and bigger headlines than did simple opining on policy issues. And this recognition sometimes led him to commandeer, wholly after the fact, regulatory decisions made in the bureaucratic trenches, with little prior or subsequent White House interest or involvement.” (footnotes omitted)).

See Presidential Remarks on Signals Intelligence, supra note 42. Speeches by the President and his senior staff have become an especially important means of shaping national security law and policy in the Obama Administration. See KENNETH ANDERSON & BENJAMIN WITTES, SPEAKING THE LAW: THE OBAMA ADMINISTRATION'S ADDRESSES ON NATIONAL SECURITY LAW 6-8 (2015); cf. RICHARD E. NEUSTADT, PRESIDENTIAL POWER 10 (1960) (“Presidential power is the power to persuade.”).

Presidential Remarks on Signals Intelligence, supra note 42.


See John Ferejohn, Accountability and Authority: Toward a Theory of Political Accountability, in DEMOCRACY, ACCOUNTABILITY, AND REPRESENTATION 131, 149 (Adam Przeworski et al. eds., 1999).

See supra notes 102-105 and accompanying text.

See generally DEMOCRACY, ACCOUNTABILITY, AND REPRESENTATION, supra note 269.

Cf. Baker, supra note 15, at 202-03 (noting that the President’s oversight responsibilities are tied to the fact that “[a]s a practical matter, he … controls access to classified information”).

See Note, Keeping Secrets: Congress, the Courts, and National Security Information, 103 HARV. L. REV. 906, 906 (1990) (“Historically, the executive branch alone has decided how to balance the need for secrecy against the need for openness in foreign affairs.”).


See Kibbe, supra note 236, at 25.


DETENTION AND INTERROGATION REPORT, supra note 63.


Jide Nzelibe, The Fable of the Nationalist President and the Parochial Congress, 53 UCLA L. REV. 1217, 1217 (2006); see also Jane Mansbridge, Representation Revisited: Introduction to the Case Against Electoral Accountability, DEMOCRACY & SOCY, Fall 2004, at 1, 12-13.

See Jack M. Beermann, Congressional Administration, 43 SAN DIEGO L. REV. 61, 143 (2006) (“Presidential supervision without effective congressional oversight is more of a threat to democratic values than congressional oversight because it can
occur privately, and the President may have been elected for reasons completely unrelated to the particular regulatory issues involved.”).

283 See Bagley & Revesz, supra note 140, at 1309 (noting OIRA’s “long and well-documented history of secrecy”).

284 See U.S. CONST. amend. IV (“The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated ....”); MacWade v. Kelly, 460 F.3d 260, 269 (2d Cir. 2006) (holding that the reasonableness of a search in the “special needs” context turns on the “efficacy of the search in advancing the government interest,” among other factors); Brief for Plaintiffs-Appellants at 2, 28, ACLU v. Clapper, 785 F.3d 787 (2d Cir. 2015) (No. 14-42) (resting their claim that the bulk collection of telephone records, under Section 215 of the Patriot Act, violates the Fourth Amendment partially upon the fact that the PRG and the PCLOB have questioned the effectiveness of the program).

285 Another approach to thinking about presidential intelligence as a vector for rights protection emphasizes the tendency on the part of centralized reviewers to be less zealous in their regulatory outlook than officials serving in agencies. If we imagine that the average intelligence officer pays less heed to the costs of his zealouleness, including costs measured in harm to privacy, than a White House overseer, then a system of presidential intelligence ought to yield greater rights protection as compared with the prior baseline of greater agency autonomy. See generally Ryan Bubb & Patrick L. Warren, Optimal Agency Bias and Regulatory Review, 43 J. LEGAL STUD. 95 (2014).

286 See Kris, supra note 190, at 292.

287 Ashley Deeks has written of the capacity of “foreign leaders, citizens, corporations, and peer intelligence services” to serve as checks on American intelligence, both directly and indirectly (by stimulating American actors to play a checking role). See Ashley Deeks, Checks and Balances from Abroad, U. CHI. L. REV. (forthcoming 2016) (manuscript at 1) (on file with the Harvard Law School Library).

288 PPD-28, supra note 43.

289 Id.

290 Henry Peck, Pull Back to Reveal: Henry Peck Interviews Ben Wizner, GUERNICA (Oct. 1, 2014), https://www.guernicamag.com/interviews/pull-back-to-reveal[https://perma.cc/UB6V-2VUZ]. As Wizner went on to say, “these tech companies, which are amassing some of the biggest fortunes in the history of the world, are among the few entities that have the power and the clout and the standing to really take on the security state.” Id.

291 See, e.g., Drumond, supra note 156. This dynamic supplies a good illustration of Professor Jon Michaels’s emphasis on translating the core insights of separation of powers to changed political and economic dynamics. See Michaels, supra note 29, at 520-21.


293 Lobbying the White House (and Congress) is not the technology firms’ sole avenue for promoting privacy. They have also gravitated toward more robust European-style privacy protections in their own services. Indeed, thanks to pressure on the companies from overseas regulators and users, these standards “have become the default privacy settings for the world.” See Mark Scott, Where Tech Giants Protect Privacy, N.Y. TIMES (Dec. 13, 2014), http://www.nytimes.com/2014/12/14/sunday-review/where-tech-giants-protect-privacy.html (quoting a former Irish data-protection official).


surveillance [http://perma.cc/N4XV-S3SU] (“How should the U.S. government do a better job of taking intelligence about known prior attacks and using that intelligence proactively to stop future ones? And that in turn will require a conversation about whether, how, and how deeply the NSA and related government agencies should be in the domestic network--not for purposes of catching Islamist terrorists, but rather for purposes of protecting our networks from other adversaries.”).

296 Presidential Remarks on Signals Intelligence, supra note 42 (“[N]othing in that initial review, and nothing that I have learned since, indicated that our intelligence community has sought to violate the law or is cavalier about the civil liberties of their fellow citizens.”). Even the ACLU’s Jaffer has conceded in respect to contemporary American electronic surveillance that “[t]he scandal is what Congress has made legal.” Glenn Greenwald & Murtaza Hussain, Meet the Muslim-American Leaders the FBI and NSA Have Been Spying On, THE INTERCEPT (July 9, 2014, 12:01 AM), https://firstlook.org/theintercept/2014/07/09/under-surveillance.


299 The PCLOB, for example, concluded that “[b]ased on the information provided to the Board, including classified briefings and documentation, we have not identified a single instance involving a threat to the United States in which the program made a concrete difference in the outcome of a counterterrorism investigation.” 215 PCLOB REPORT, supra note 244, at 11.


303 We also know that there was a period of time when the FISC expressed its concerns about the NSA’s implementation of the program, during which the court assumed the responsibility, in effect, for checking the NSA’s compliance work. See Jennifer Granick, New FISC Pen Register Opinion: It's Just a Matter of Time Before Somebody Gets Hurt, JUST SECURITY (Nov. 21, 2013, 7:06 PM), http://justsecurity.org/3576/fisc-pen-register-opinion-its-matter-time-hurt [http://perma.cc/U9KW-V4CE].


305 Id. tit. I, § 101. The legislation was prompted by the sunsetting of a number of Patriot Act authorities on May 31, 2015. Among other things, the new law includes the designation by the FISC of a panel of amici curiae to serve in cases raising novel or significant legal issues. Id. tit. IV, § 401.
306 See infra pp. 704-06.

307 See PPD-28, supra note 43 (extending certain privacy protections to non-U.S. persons). Given this attention, it is somewhat exaggerated for Kenneth Roth to argue that “Obama has not addressed another troubling aspect of US electronic surveillance—the view that foreign citizens outside the United States have no right to privacy even in the content of their communications.” See Roth, supra note 240.

308 See, e.g., Jody Freeman & Adrian Vermeule, Massachusetts v. EPA: From Politics to Expertise, 2007 SUP. CT. REV. 51, 87 (“This approach hearkens back to an older, pre-Chevron vision of administrative law in which independence and expertise are seen as opposed to, rather than defined by, political accountability, and in which political influence over agencies by the White House is seen as a problem rather than a solution.”); The Supreme Court, 2006 Term—Leading Cases, 121 HARV. L. REV. 185, 420 (2007).

309 Kagan, supra note 1, at 2353-54; see also Peter L. Strauss, Foreword, Overseer, or “The Decider”? The President in Administrative Law, 75 GEO. WASH. L. REV. 696, 752 (2007) (“While Chevron sensibly accepts the President's political role as mediating the difficulties of focused bureaucratic expertise, it does not purport to displace reliance on the latter. Indeed, the structure of judicial review of administrative action depends, top to bottom, on the presumption that the matter being reviewed is in some respects the product of an expert, not merely a political judgment.”).

310 See, e.g., Barron, supra note 98, at 1150 (“There is a great deal of science on the issue of global warming, obviously. But the fact that there is a scientific consensus on the role that human activity plays in causing climate change hardly answers the policy question of what should be done in response. Thus, an embrace of scientific expertise alone cannot resolve the hardest policy questions in this area any more than is usually the case.” (footnote omitted)).

311 The issue was politically salient during the Bush Administration, with concerns that the White House “embraced bad politics over science in certain areas ... and ignored science for bad political reasons in other areas.” RENA STEINZOR & SIDNEY SHAPIRO, THE PEOPLE'S AGENTS AND THE BATTLE TO PROTECT THE AMERICAN PUBLIC 132 (2010). President Obama's White House has not been immune to this tendency. See Mendelson, supra note 223, at 1143 (discussing the Obama Administration's failure to acknowledge the political considerations that informed its decision to prohibit women under 18 from buying so-called Plan B contraceptives over the counter).

312 Freeman & Vermeule, supra note 308, at 52. Their reading of the case is not naïve, in that it recognizes that “[a]ll administrations exert political pressure on their executive agencies,” and that “it is inevitable that political considerations will come into play in executive agencies headed by political appointees who are accountable to the President.” Id. at 108-09. For the sake of precision it is important to note that the main argument in the case was that greenhouse gases are not “air pollutants” for the purposes of the Clean Air Act. See Massachusetts v. EPA, 549 U.S. 497, 528 (2007). A clearer climate-denying claim would have been that they don’t “endanger public health,” a statutory claim that was not before the Court.

313 Scholars have, of late, attempted to flesh out what has been a theoretically impoverished conversation on the meaning of politicization. See, e.g., JOSHUA ROYNER, FIXING THE FACTS: NATIONAL SECURITY AND THE POLITICS OF INTELLIGENCE 36-48 (2011). Professor Richard Betts has questioned the reflexively critical posture that commentators bring to discussions of politicization. See RICHARD K. BETTS, ENEMIES OF INTELLIGENCE: KNOWLEDGE AND POWER IN AMERICAN NATIONAL SECURITY 74 (2007).

314 AGRELL & TREVERTON, supra note 51, at 162.

315 SHERMAN KENT, STRATEGIC INTELLIGENCE FOR AMERICAN WORLD POLICY 200 (1949).

316 See BETTS, supra note 313, at 76.

See Paul R. Pillar, *Intelligence, Policy, and the War in Iraq*, FOREIGN AFF., Mar.-Apr. 2006, https://foreignaffairs.com/articles/iraq/2006-03-01/intelligence-policy-and-war-iraq (“That the administration arrived at so different a policy solution indicates that its decision to topple Saddam was driven by other factors--namely, the desire to shake up the sclerotic power structures of the Middle East and hasten the spread of more liberal politics and economics in the region.”).


See AGRELL & TREVERTON, supra note 51, at 51. Some intelligence analysis, notably the State Department’s Bureau of Intelligence and Research (INR), continues to be organized this way. See Bureau of Intelligence and Research, U.S. DEPT OF STATE, http://www.state.gov/s/inr [http://perma.cc/3EC8-QZTF].

See, e.g., Melvin A. Goodman, Opinion, *Separate the C.I.A.’s Intelligence and Operations*, N.Y. TIMES (Dec. 21, 2014), http://www.nytimes.com/roomfordebate/2014/12/21/do-we-need-the-cia/separate-the-cias-intelligence-and-operations (maintaining that CIA centers that fuse operational and analytic functions “undermine[] the ability of analysts to provide objective analysis”).

As Michael Hayden once put it, “[i]f it were a fact, it wouldn’t be intelligence.” BOB WOODWARD, PLAN OF ATTACK 219 (2004).


See supra notes 76-81 and accompanying text.

See Kagan, supra note 1, at 2356-57; cf. Lisa Schultz Bressman & Robert B. Thompson, *The Future of Agency Independence*, 63 VAND. L. REV. 599, 634-35 (2010) (“Bernanke recognize[d] that for decisions so profoundly national in scope, the combination of politics and expertise is more powerful than expertise alone. All else equal, the President is likely to have information that is relevant to generating sound policy on market stability and to mobilizing the necessary political will to achieve the results.”).


See ROVNER, supra note 313, at 203-04.

See infra section IV.A, pp. 706-12.

See, e.g., Kagan, supra note 1, at 2347 (“Congress, of course, always faces disincentives and constraints in its oversight capacity .... Because Congress rarely is held accountable for agency decisions, its interest in overseeing much administrative action is uncertain; and because Congress’s most potent tools of oversight require collective action (and presidential agreement), its capacity to control agency discretion is restricted.”).

See David J. Arkush, *Direct Republicanism in the Administrative Process*, 81 GEO. WASH. L. REV. 1458, 1478-79 (2013) (“[T]he notion of ‘congressional’ oversight, in the sense of the whole Congress watching over regulators, is rarely more than a metaphor. Legislative supervision typically takes the form of oversight by a small number of individuals in Congress,
usually the heads of relevant committees or, more specifically, their staffs, some of whom may be as removed from electoral accountability as agency officials.” (footnotes omitted)).

332 See id. at 1479 (“[C]riticisms point to a circularity in aspirations for congressional oversight: Congress delegates broad authority to administrative agencies because it is unwilling or unfit to make all of the decisions required in various policy areas. If Congress were willing and able to evaluate agency performance on the relevant matters, then it need not have delegated the authority in the first place.”).

333 See Moe & Wilson, supra note 91, at 11-12.

334 See Brian D. Feinstein, Congressional Government Rebooted: Randomized Committee Assignments and Legislative Capacity, 7 Harv. L. & Pol’y Rev. 139, 160 (2013) (“Despite the theoretical importance and demonstrated efficacy of oversight, Congress appears relatively uninterested in performing its oversight function. Oversight-focused subcommittees tend to be disproportionately populated by less powerful legislators, with senior legislators, party leaders, and full committee chairs and ranking members rarely serving on subcommittees devoted to oversight and investigatory work.”); Douglas Kriner, Can Enhanced Oversight Repair “the Broken Branch”?, 89 B.U. L. Rev. 765, 792 (2009) (“[R]forms do little to address the underlying problem of variable congressional motivation to oversee the executive in the first place.”).

335 See Strauss, supra note 309, at 759-60 (“Congress can, to be sure, give the President decisional authority, and it has sometimes done so. In limited contexts-- foreign relations, military affairs, coordination of arguably conflicting mandates--the argument for inherent presidential decisional authority is stronger. But in the ordinary world of domestic administration, where Congress has delegated responsibilities to a particular governmental actor it has created, that delegation is a part of the law whose faithful execution the President is to assure. Oversight, and not decision, is his responsibility.” (footnote omitted)).


338 Id. at 559.

339 See ZEGART, supra note 32, at 74-75.


341 See Anne Joseph O'Connell, The Architecture of Smart Intelligence: Structuring and Overseeing Agencies in the Post-9/11 World, 94 CALIF. L. Rev. 1655, 1671 (2006) (“While Congress and the Administration have made at least some serious efforts to reorganize the intelligence community, Congress has made little effort to reorganize its overlapping committee oversight of the intelligence community.”).

342 See, e.g., NAT'L COMM’N ON TERRORIST ATTACKS UPON THE U.S., THE 9/11 COMMISSION REPORT 103 (2004) (discussing fragmented nature of congressional appropriations and oversight of intelligence organizations); see also id. at 419-21 (calling congressional oversight of intelligence community “dysfunctional” and recommending each house of Congress create a single committee responsible for intelligence oversight to prevent attacks in the future).

343 See Johnson, supra note 141.

Zegart, supra note 31, at 215.


L. Britt Snider, Congressional Oversight of Intelligence: Some Reflections on the Last 25 Years 10, (unpublished manuscript) https://web.law.duke.edu/lens/downloads/snider.pdf [https://perma.cc/NCA5-LG69] (“For all of the situations in which the oversight committees--even today--might find themselves in the dark, they are, for the most part, aware of what the intelligence agencies are doing, however sensitive those activities might be.”).


See Eli Lake, Congress Scouring Every U.S. Spy Program, DAILY BEAST (Oct. 10, 2014, 5:45 AM), http://www.thedailybeast.com/articles/2014/10/10/congress-scouring-every-u-s-spy-program.html [http://perma.cc/L97Q-GNBE] (quoting a senior committee staff member describing the investigation's unprecedented scope as encompassing "[a]ll the programs through which the intelligence community collects intelligence. Human intelligence, signal intelligence, open source. It is all subject to the review").


Id. at 2363 (quoting Cynthia R. Farina, Undoing the New Deal Through the New Presidentialism, 22 HARV. J.L. & PUB. POLY 227, 235 (1998)).

See, e.g., Samuel Issacharoff, Political Safeguards in Democracies at War, 29 OXFORD J. LEGAL STUD. 189, 207 (2009) ("[I]t is a matter of profound dishonor to our constitutional system that, prior to the 2006 elections, not once did Congress, under the control of the Republican Party, hold meaningful hearings over the conduct of the Iraq War by a Republican President.").

See id.


See DETENTION AND INTERROGATION REPORT, supra note 63.

Indeed, the current Republican Chairman, Senator Richard Burr “has demanded that the Obama administration return every copy of the report.” See Mark Mazzetti & Matt Apuzzo, Classified Report on the C.I.A's Secret Prisons Is Caught in Limbo, N.Y. TIMES (Nov. 9, 2015), http://www.nytimes.com/2015/11/10/us/politics/classified-report-on-the-cias-secret-prisons-is-caught-in-limbo.html. Professor Aziz Huq argues that when it comes to congressional intelligence oversight, “failure has been fairly constant regardless of whether we have divided government or the same party in the White House and in command on Capitol Hill.” Huq, supra note 5.


This select group of elected officials, consisting of the chairmen and ranking minority members of the congressional intelligence committees, the speaker and minority leader of the House, and the majority and minority leaders of the Senate, has privileged access to certain high-level intelligence matters. See 50 U.S.C. § 3093(c)(2) (Supp. 1 2013).


*Id.*


This explanation is predicated on the classic Madisonian conception that, as Congress put it in the counterterrorism context, a “shift of power and authority to the Government calls for an enhanced system of checks and balances.” 42 U.S.C. § 2000ee(b) (2) (2012).


See Rascoff, *supra* note 60, at 586.

The provision for amici curiae at the FISC as part of the recently enacted USA Freedom Act may go some way toward addressing the limitations of the preexisting system. See *supra* note 127.

See *supra* note 65 and accompanying text.

See, e.g., Clapper v. Amnesty Int'l USA, 133 S. Ct. 1138, 1143 (2013) (holding that plaintiffs failed to demonstrate that their claimed injuries were fairly traceable to the government's surveillance activities at issue).


ACLU v. Clapper, 785 F.3d 787, 821 (2d Cir. 2015) (holding that the telephone metadata collection program exceeded the scope of what Congress authorized in section 215).

Clapper, 133 S. Ct. at 1143.


382 See Issacharoff, supra note 354, at 206 (“However much the burdens of wartime democracy must rest on the political branches, there remains the need for a fuller rendition of the role of the judiciary.”).


386 Cf. ERWIN, supra note 61.


388 ERWIN, supra note 61, at 6. Further, the finding “may not authorize any action intended to influence” domestic affairs and “may not authorize any action which violates the Constitution ... or any statutes of the United States.” Id. (quoting 50 U.S.C. § 413b (recodified as 50 U.S.C. § 3093 (Supp. 1 2013))); see also Benjamin Powell, Secret Operations: Covert Action and Military Activities, in NATIONAL SECURITY LAW IN THE NEWS: A GUIDE FOR JOURNALISTS, SCHOLARS, AND POLICYMAKERS 123, 128-29 (Paul Rosenzweig et al. eds., 2012).

389 See Daugherty, supra note 387, at 75.

390 WILLIAM J. DAUGHERTY, EXECUTIVE SECRETS: COVERT ACTION AND THE PRESIDENCY 93-94 (2004) (noting that this procedure “was the practical death of ‘plausible deniability,’” id. at 94); LESTER, supra note 58, at 138 (“The finding is a critical legal document and is usually briefed to the president in person.”).

391 See Marshall Silverberg, The Separation of Powers and Control of the CIA’s Covert Operations, 68 TEX. L. REV. 575, 602 (1990) (“The ... debate [in 1990] ... focused on the time period within which the President must inform Congress of a covert action.”). The congressional notification requirement is “not a precondition to carrying out a covert action” and its implementation may be delayed “until after the operation has commenced or occurred” in some unique circumstances. Powell, supra note 388, at 129. However, in these circumstances, the President is still required to report to the committees “in a timely fashion and ... provide a statement of the reasons for not giving prior notice.” Id. (omission in original) (quoting 50 U.S.C. § 413b(c)(3) (recodified as 50 U.S.C. § 3093(c)(3) (Supp. 1 2013))).

392 See Daugherty, supra note 387, at 63. Former CIA Acting General Counsel John Rizzo relates that the idea for the Memorandum of Notification originated with White House Counsel Lloyd Cutler in 1980. Cutler was concerned that President Carter was signing too many Findings, and so asked CIA lawyers to “create a different name for a Finding that simply expands or otherwise changes the scope of a preexisting Finding.” See RIZZO, supra note 101, at 75.


Additionally, if the President “determines it is essential to limit access to the finding to meet extraordinary circumstances affecting vital interests of the United States,” reporting can be limited to the Gang of Eight. 50 U.S.C. § 3093(c)(2) (Supp. I 2013).


§ 503, 105 Stat. at 442-44.

See Daugherty, supra note 387, at 68-73 (describing the process under President Reagan and President Clinton).

Id. at 68.

LOCH K. JOHNSON, AMERICA'S SECRET POWER: THE CIA IN A DEMOCRATIC SOCIETY 21 (1989); see also id. at 20 (quoting former Secretary of Defense and National Security Act of 1947 author Clark Clifford as saying “the guiding criterion ... should be the test as to whether or not a certain covert project truly affects our national security” (emphasis added)).


John Prados, The Continuing Quandary of Covert Operations, 5 J. NAT'L SECURITY L. & POL'Y 359, 362-68 (2012) (arguing that the CIA's lack of “detailed knowledge that might enable it to choose among alternatives,” id. at 365, as well as its inability to anticipate potential “blowback,” id. at 363, call into question the “finding” methodology).

See Johnson, supra note 62, at 54-55 (positing that general, broad wording of findings such as “‘to fight global terrorism’ ... could be interpreted to mean almost anything, with carte blanche [for the CIA] to carry out operations anywhere, anytime”); Austin Long, US Covert Action Success and Intelligence Policy 3 (Naval War College Conference Paper) (Mar. 25, 2013), https://www.usnwc.edu/Academics/Faculty/Derek-Reveron/Workshops/Intelligence,-National-Security-and-War/Documents/Long.aspx (last visited Nov. 22, 2015) (“The crafting of the finding's language can either facilitate or hamper tactical improvisation in the field, and is one indicator of the level of trust the President has in covert action.”); Pam Benson, What's Allowed by a “Presidential Finding”? CNN (Mar. 31, 2011, 9:24 PM), http://www.cnn.com/2011/POLITICS/03/31/ libya.presidential.finding (quoting a former intelligence official for the proposition that findings are “written in a way that is ‘general enough to allow flexibility, but specific enough to know legally what you can do’”).

DETENTION AND INTERROGATION REPORT, supra note 63, Executive Summary at 11.

Id., Findings and Conclusions at 9.


See MARK MAZZETTI, THE WAY OF THE KNIFE: THE CIA, A SECRET ARMY, AND A WAR AT THE ENDS OF THE EARTH 217 (2013) (“Obama's desire to manage aspects of the targeted-killing program directly from the White House gave Brennan a role unique in the history of American government: one part executioner, one part chief confessor to the president, one part public spokesman sent out to justify the Obama doctrine of killing off America's enemies in remote parts of the world.”). The President himself made a connection between drone strikes and surveillance practices, observing in his January 2014 speech that “after an extended review of our use of drones in the fight against terrorist networks, I believed a fresh examination of our surveillance programs was a necessary next step in our effort to get off the open-ended war footing that we've maintained since 9/11.” Presidential Remarks on Signals Intelligence, supra note 42.
It may seem strange to use covert action—which, by statute, refers to certain activities undertaken by the U.S. government as to which “the role of the United States Government will not be apparent or acknowledged publicly,” 50 U.S.C § 3093(e) (2014)—as a template for presidential intelligence under conditions of ever-greater visibility. But, as a practical matter, the contrast need not be that great. The bin Laden raid counted as covert action even though the White House took public credit for it hours after it transpired. See Remarks on the Death of Al Qaida Terrorist Organization Leader Usama bin Laden, 1 PUB. PAPERS 480 (May 1, 2011).

PPD-28, supra note 43.

See PRG, supra note 65, at 167.

Omand, supra note 132, at 624.


In the (admittedly, quite different) context of OIRA review, only “major” rules qualify for presidential cost-benefit analysis. Exec. Order No. 12,291, 3 C.F.R. 200 (1982). One way for a rule to be deemed “major” is if it has “an annual effect on the economy of $100 million or more.” Id. Another is if it will likely result in a “major increase in costs” for “consumers [or] individual industries.” Id.

The President's Review Group proposed a number of potential criteria for determining what sort of intelligence activities senior policymakers should review, including: (1) the means of collection; (2) the specific people being monitored; (3) the country in which the collection takes place; (4) any international meeting where the collection takes place; or (5) some combination of these factors. See PRG, supra note 65, at 168-69.

Cf. Bressman & Vandenberg, supra note 223, at 97-98 (“If regulatory review is extended to include scientific review, OIRA (or other White House offices) should acquire additional scientific expertise.”).


Cf. PRG, supra note 65, at 167 (“A small staff of policy and intelligence professionals should review intelligence collection for sensitive activities on an ongoing basis throughout the year and advise the National Security Council Deputies and Principals when they believe that an unscheduled review by them may be warranted.”).


Moe & Wilson, supra note 91, at 18. Professors Robert Durant and William Resh put forward a number of reasons that a politicization strategy “may not only be difficult but may actually be counterproductive.” Robert F. Durant & William G. Resh, “Presidentializing” the Bureaucracy, in THE OXFORD HANDBOOK OF AMERICAN BUREAUCRACY, supra note 31, at 545, 553. Of the reasons that they summon, (at least) one is particularly applicable to intelligence, namely the concern that short-term nonexpert appointees would become beholden to the expert professional bureaucracy. Id. But in a world of greater uses of “intelligence” in the private sector, it may be increasingly plausible to draw on knowledgeable outsiders to assume leadership roles within the intelligence bureaucracy.

Cf. Moe, supra note 46, at 150 (“The legacy of the past discourages comprehensive reform efforts—but, precisely because it does, it magnifies the president's incentives to pour effort into minor but feasible changes by making maximum use of the structures and resources closest to him .....”).
A lack of presidential control does not characterize all security agencies. As far back as President Franklin Roosevelt's administration, Presidents have sought to control the national security bureaucracy through a combination of centralization and politicization. See MARIANO-FLORENTINO CUÉLLAR, GOVERNING SECURITY: THE HIDDEN ORIGINS OF AMERICAN SECURITY AGENCIES 88 (2013) (“[T]he Roosevelt administration ... sought substitutes for the White House staff increases in the form of new layers of political appointees to oversee existing bureaus ...”).


See Baker & Mazzetti, supra note 47. Unlike Casey, Brennan is a career intelligence officer who previously spent twenty-five years at the CIA. Different CIA directors have adopted different postures toward the President. Writing in Time Magazine about the close personal tie between President Obama and Brennan, former spy Robert Baer observed that “[t]he last CIA director with a close personal relationship with his President was Reagan's CIA director Bill Casey” who, Baer goes on to say, “played an important role in shaping Reagan's foreign policy.” Robert B. Baer, What Awaits John Brennan at the CIA, TIME (Jan. 9, 2013), http://swampland.time.com/2013/01/09/what-awaits-john-brennan-at-the-cia [http://perma.cc/FA2L-BWF3]. Notably, Casey was the first (and perhaps also the last) CIA Director to “take a place at the White House table as a fully participating Cabinet member.” See Pace, supra note 423.


See Robertson, supra note 16, at 249.

See CHRISTOPHER M. DAVIS & JERRY W. MANSFIELD, CONG. RESEARCH SERV., RL30959, PRESIDENTIAL APPOINTEE POSITIONS REQUIRING SENATE CONFIRMATION AND COMMITTEES HANDLING NOMINATIONS 38 (2013); PLUM BOOK, supra note 112.

See PLUM BOOK, supra note 112, at 141.


See Kirti Datla & Richard L. Revesz, Deconstructing Independent Agencies (and Executive Agencies), 98 CORNELL L. REV. 769, 791 (2013) (“[T]he decision to remove an officer before the end of a specified term imposes at least some costs on a President.”).

policymakers take a longer-term perspective when making their decisions. Policymakers in an independent central bank, with a mandate to achieve the best possible economic outcomes in the longer term, are best able to take such a perspective.”). This rationale for independence is not particularly applicable to the intelligence domain.


436 Although the PRG recommended divesting the NSA Director of the responsibility to lead Cyber Command, the President ultimately decided to maintain the status quo. See Ellen Nakashima, White House to Preserve Controversial Policy on NSA, Cyber Command Leadership, WASH. POST (Dec. 12, 2013), http://www.washingtonpost.com/world/national-security/white-house-to-preserve-controversial-policy-on-nsa-cyber-command-leadership/2013/12/13/4bb56a48-6403-11e3-a373-0f9f2d1c2b61_story.html [http://perma.cc/JTV5-DYA9].


439 PLUM BOOK, supra note 112, at 198.


441 Any additional politicization of the intelligence agencies would need to take into account the rise of centralized review, for the two phenomena are necessarily interactive. See Bubb & Warren, supra note 285, at 100-01.

442 Cf. RIZZO, supra note 101, at 44-47 (discussing the growing number and prominence of lawyers in the CIA).

443 Here a comparison with financial regulation may be apt. In that domain, the most qualified potential political appointees are likely to be veterans of the financial services industry. See, e.g., Jim Tankersley, Wall Street Veteran Heads New Federal Office Tasked with Making Better Economic Forecasts, WASH. POST (Apr. 5, 2013), http://www.washingtonpost.com/business/wall-street-veteran-heads-new-federal-office-tasked-with-making-better-economic-forecasts/2013/04/05/bc9c912e-9ad6-11e2-9a79-eb5280c81e63_story.html [http://perma.cc/34DJ-S8WY].

444 Certainly caution must be exercised here, as the problematic tenure of Porter Goss at the CIA reveals. See Dana Priest & Walter Pincus, CIA Chief Seeks to Reassure Employees, WASH. POST (Nov. 16, 2004), http://www.washingtonpost.com/wp-dyn/articles/A51301-2004Nov15.html [http://perma.cc/8P4Q-7DV6] (“His critics say the director, a former CIA case officer and Republican chairman of the House intelligence committee, is purging the agency of career officers whom he incorrectly perceives as critical of Bush administration policies. In addition, Goss has over the last month put in charge several former Hill staff members who are not well regarded by senior officials because they lack managerial and operational experience, and are believed to have treated career officers disrespectfully.”). Goss lasted less than two years in the job.

445 For example, President Bush (briefly) managed through Executive Order 13,422 to require that so-called “regulatory policy officers” who play key roles in agency regulatory planning and cost-benefit analysis be presidential appointees. See CURTIS W. COPELAND, CONG. RESEARCH SERV., RL33862, CHANGES TO THE OMB REGULATORY REVIEW PROCESS BY EXECUTIVE ORDER 13,422, at CRS-6 (2007).
See Anne Joseph O'Connell, *Vacant Offices: Delays in Staffing Top Agency Positions*, 82 S. CAL. L. REV. 913, 913-14 (2009) (describing the difficulties Presidents face in filling appointed positions and pointing out that the average Senate-confirmed position was empty or filled by an acting official approximately one-quarter of the time).


See PODESTA ET AL., *supra* note 75 (considering the potential impacts of Big Data on health care, education, homeland security, and law enforcement).


129 HVL R 633
INTRODUCTION

The encryption debate over whether there should be a lawful access requirement has re-emerged in recent years with major security implications. This debate has become part of the larger discussion about balancing changing technological capabilities, national security threats, and the evolution of society's sense of privacy. Although this is an important area of legal policy, unfortunately, thus far the opposing sides of this issue have been unwilling to engage in meaningful deliberations with those on the other side. This Article thoroughly analyzes the legitimate arguments that exist on both sides with the aim of finding a reasonable solution to the encryption debate.
Law enforcement and intelligence officials fear that they may not be able to prevent terrorist attacks or conduct criminal investigations and prosecute criminal activity without access to communications. Thus, law enforcement and intelligence agencies have advocated for a lawful access requirement to mandate that companies maintain access to users' communications and data, and provide law enforcement or intelligence agencies with that access upon receipt of a lawful order. However, there are serious concerns that a mandated lawful access requirement would increase users' vulnerability to various cyber-threats, and would decrease the market share and economic viability of U.S. companies. Also, this debate over whether there should be a lawful access requirement has privacy and civil liberties implications in the United States and around the world. Ultimately, the United States should mandate a lawful access requirement, and it is technologically feasible to securely achieve this. This will enable law enforcement and intelligence agencies to continue pursuing their security missions in an effective manner, while maintaining cybersecurity and U.S. companies' market share and economic viability.

Part I provides a brief explanation of encryption, and the two types of encryption that are at the center of this debate, which are end-to-end encryption and endpoint encryption--also called device encryption. Part II discusses the history of the encryption debate. The current debate is actually not the first time that encryption has been a focal point of heated legal, policy, security, and privacy discussions. The government's access to encrypted communications has been debated since the 1970s, and first came to a head in the 1990s. This has been called the “Crypto Wars.” Part III examines the two developments that have spurred the second coming of the “Crypto Wars.” First, encryption is increasingly *67 becoming the default setting on devices. Second, service providers are offering end-to-end encryption on products and encrypting data that is stored in cloud storage systems.

Both sides of this debate present compelling arguments, and Part IV thoroughly analyzes law enforcement and intelligence concerns about encryption as well as private sector and security concerns about mandating a lawful access requirement. Law enforcement and intelligence agencies are concerned that they are “going dark” because there is an increasing number of electronic communications that they have the legal authority to intercept, but cannot feasibly do so. The private sector and security officials have expressed numerous reasons why they oppose a lawful access requirement. The most prominent among these arguments is the fear that the technological architecture that would guarantee law enforcement and intelligence agencies access would compromise user security and privacy. This Article argues that the greatest potential harm from requiring lawful access is actually not the potential for user security and privacy to be compromised; rather, the greatest potential harm is the possible decrease in the market share and economic viability of U.S. companies. This aspect of the debate has not been analyzed enough, and this Article finds that the common assumption that U.S. companies will lose market share and economic viability because of a lawful access requirement is incorrect. After a thorough examination of both sides of the debate and the alternatives that have been put forth to a lawful access requirement, this Article concludes that Congress should mandate a lawful access requirement by requiring companies to have the capacity to deliver decrypted information upon applicable legal process.

Subsequently, Part V takes on the challenge of examining how a lawful access requirement could be achieved by analyzing several current proposals. Technological innovations could be developed, users could be given greater control, or industry incentives could be altered through conditional liability protections. Ultimately, Part VI supports the development of technological innovations to achieve the lawful access requirement that this Article calls on Congress to enact.

I. WHAT IS ENCRYPTION?

Encryption is the process of encoding data or information such that only those who are authorized by the creator of the information are able to access the data or information. ¹ Those who are not authorized by the *68 creator of the data
or information to have access are prevented access to encrypted data or information. Even if a third party without authorization intercepts the data or information, encrypted data or information will appear unreadable.

The types of encryption at the heart of this debate are end-to-end encryption and endpoint encryption--also called device encryption. End-to-end encryption is the encryption of messages in transit such that only the original sender and intended recipient hold the keys to decrypt the communication. The message in transit can therefore only be read by the original sender and intended recipient. This type of encryption is important for protecting data in motion. Endpoint encryption is when the keys only exist on locked devices, which prevents the contents of the device from being read by anyone who does not possess the keys. This type of encryption is important for protecting stored information--data at rest.

II. THE “CRYPTO WARS”

The current encryption debate is not the first time this subject has been analyzed. In fact, government access to encrypted communications has been debated since the 1970s. This ongoing debate has been called the “Crypto Wars.” The National Security Agency (NSA) and other intelligence agencies advocated for the export control and restriction of strong encryption. They sought to limit the ability of technology companies to export products with strong encryption and to inhibit researchers from publishing cryptographic research. Similar to today, intelligence agencies worried that encryption could be used by adversaries in a manner that would make it more difficult to conduct signals intelligence (SIGINT), which would threaten U.S. national security.

The “Crypto Wars” came to a head in the 1990s as encryption technologies became more widespread. Relatively low-cost computers gained enough processing power in the 1990s “to be capable of encrypting data with a level of security that would make it nearly impossible for the government to break.” Law enforcement agencies believed the spread of this technology would adversely affect their investigations and joined the effort to limit encryption. Furthermore, the Federal Bureau of Investigation (FBI) was concerned with telephone companies' move from copper wires to fiber optics in the 1990s. The move to fiber optics made traditional telephone wiretaps less useful. In 1992, the FBI warned that by 1995 at most only 60% of wiretaps would be useful, and at worst wiretaps would be completely useless.

On the other hand, the technology industry, technologists, and civil libertarians opposed the government's hindrance of cryptology. The technology industry and technologists believed widespread encryption was beneficial because the security of critical infrastructure and financial markets were increasingly becoming dependent on strong cybersecurity. Civil libertarians believed the government's export control restrictions violated cryptographers' First Amendment rights and that encryption enhanced privacy.

The government sought to combat encryption through continued export control and key escrow, in which the government or a neutral third party would hold the keys in escrow. In 1993, the NSA developed the Clipper Chip as a key escrow system. The Clipper Chip was designed for phones to provide encryption while also producing an encryption key that would be available to the NSA. Although implementation of the Clipper Chip was voluntary, it was the federal standard, which made the government optimistic that it would become widespread because industries that dealt with the federal government would need to adopt the system. The Clipper Chip was met with a great deal of backlash, though. People feared that the chip would compromise their privacy. Also, shortly after the first Clipper
Chip products entered the market, a cryptographer, Matt Blaze, discovered and publicly disclosed vulnerabilities in the Clipper Chip. This led the government to cease its promotion of the Clipper Chip.

Although the government's key escrow effort failed, law enforcement was able to secure the passage of the Communications Assistance to Law Enforcement Act (CALEA) in 1994. CALEA addressed the concern that the shift from copper wires to fiber optics made traditional wiretaps less useful by requiring telecommunications carriers to ensure their networks could be wiretapped pursuant to a court order. CALEA was intended to preserve the ability of law enforcement officials to conduct electronic surveillance effectively and efficiently, despite the deployment of new digital technologies and wireless services by the telecommunications industry. CALEA requires telecommunications carriers to modify their equipment, facilities, and services to ensure that they are able to comply with authorized electronic surveillance.

Thus, the government was at least able to ensure that law enforcement could continue to conduct electronic surveillance regardless of the type of telecommunications system used. However, CALEA did not address encryption. The statute did not authorize law enforcement “to require any specific design of equipment, facilities, services, features, or system configurations to be adopted” or “to prohibit the adoption of any equipment, facility, service, or feature.” In addition, CALEA included the caveat that “[a] telecommunications carrier shall not be responsible for decrypting, or ensuring the government's ability to decrypt, any communication encrypted by a subscriber or customer, unless the encryption was provided by the carrier and the carrier possesses the information necessary to decrypt the communication.” Further, CALEA's focus was on telecommunications carriers, not Internet-based communications services, which were still in an early stage of development at the time of CALEA's passage. Although CALEA has been interpreted by the Federal Communications Commission (FCC) to expand the statute's reach to “facilities-based broadband Internet access and Voice over Internet Protocol (VoIP) services that are fully interconnected with the public switched telephone network,” CALEA does not cover “popular Internet-based communications services such as email, Internet messaging, social networking sites, or peer-to-peer services.”

Another way the government tried to limit the spread of encryption was that encryption software and related technologies continued to remain on the U.S. Munitions List through the early 1990s, which meant that encryption was treated as a defense-related technology and its export was heavily regulated by the Department of State. In 1996, President Bill Clinton removed commercial encryption products from the U.S. Munitions List but continued to mandate that only “export-grade” encryption could be shipped overseas rather than the stronger encryption that was being developed. In 1999, the Ninth Circuit Court of Appeals ruled that the government's export controls, which restricted the publication of source code, violated cryptographers' First Amendment rights to free speech in *Bernstein*. This, along with continued backlash from industry and advocacy groups, as well as international trade considerations resulted in the government relaxing export controls.

Despite losing the “Crypto Wars” regarding encryption, the government's fears of the downsides of encryption did not come to fruition. Instead, encryption went largely unused by the general public. Encryption's complexity made it too difficult to be used in a widespread manner. Also, the FBI received greater resources to enhance its technical capabilities, and the government worked with industry to develop a partnership to enable lawful intercepts to be conducted.
III. THE DEBATE RE-EMERGES

Two recent developments have spurred the re-emergence of this debate. First, encryption is increasingly becoming the default setting on devices. Data was formerly stored on devices in an unencrypted form unless the user took affirmative action to use encryption. Now, however, more devices will encrypt data by default unless the user takes affirmative action to turn this function off and store data in an unencrypted form. Thus, the burden of action formerly favored not using encryption, whereas now the burden of action will favor using encryption. This will greatly increase the prevalence of endpoint encryption. Apple, in particular, has been a leader in promoting default endpoint encryption. In 2014, Apple announced it would include default encryption of its devices that use the iOS 8 mobile operating system. Google followed suit by making encryption the default on its Android operating system.

Second, service providers are offering end-to-end encryption on products and encrypting data that is stored in cloud storage systems. These products encrypt data, information, and communications in such a way that the service provider does not have the technical capability to decrypt the information. Therefore, these providers cannot respond to lawful process because they do not possess the information the government is requesting. In 2014, WhatsApp, an instant messaging service on smartphones that is now owned by Facebook, announced it would use end-to-end encryption. Other applications that have also implemented end-to-end encryption have become popular recently, too.

IV. THE CURRENT DEBATE

A. Law Enforcement and Intelligence Concerns

The FBI has led the government's involvement in the current debate since 2010. The FBI is concerned that it is “going dark” because there is an increasing number of electronic communications that the FBI has the legal authority to intercept, but cannot feasibly do so. Reports have indicated that encryption and other technological means, like proxy servers, can conceal information from lawful electronic surveillance. Intelligence agencies, especially the NSA and Central Intelligence Agency (CIA), also face difficulties in fulfilling their missions of gathering intelligence because of encryption. The NSA and CIA have greater resources to combat this problem than the FBI, though. Furthermore, the “going dark” problem is most acute for state and local law enforcement agencies that have fewer resources than federal law enforcement. Nonetheless, there is concern across law enforcement and intelligence agencies about “going dark,” and these agencies at all levels of government “would benefit if technological architectures did not present a barrier to investigations.”

Ultimately, law enforcement and intelligence agencies fear that they will not be able to prevent terrorist attacks, investigate crimes, and prosecute criminal activity without access to communications. FBI Director James Comey, who has been very vocal in describing the “going dark” problem has stated,

[unfortunately, the law hasn't kept pace with technology, and this disconnect has created a significant public safety problem ... Those charged with protecting our people aren't always able to access the evidence we need to prosecute crime and prevent terrorism even with lawful authority. We have the legal authority
to intercept and access communications and information pursuant to court order, but we often lack the technical ability to do so. 53

Without a lawful access requirement, there will be crimes that go unsolved that otherwise may have been solvable, and criminals will not be brought to justice.

Currently, there is a double homicide case in Baton Rouge, Louisiana that remains unsolved where the information on the victim's phone is encrypted, and the phone cannot be opened without the password. 54 The *75 victim was a pregnant woman who was shot and killed, and whose unborn baby died. 55 The victim's daughter was at the apartment and heard her mother open the door and speak with a person at the door before the daughter heard gunshots and hid in the bathroom. 56 There was no sign of forced entry at any other apartment in the complex, and the shooter did not enter the victim's apartment. 57 Critically, the victim kept a detailed diary on her phone, which could be useful for law enforcement investigators. 58 Investigators suspect the victim may have information regarding the shooter in her diary because she was willing to open the door for, and speak with, the shooter. 59 However, law enforcement cannot obtain access to the phone because it is locked, and the unencrypted data stored on iCloud stopped backing up several months prior to the murder. 60 Thus, the heinous crime remains unsolved and law enforcement does not yet even have a suspect. 61

Also, the FBI has revealed that one of the terrorists who attempted to attack Garland, Texas exchanged 109 encrypted messages with a terrorist overseas. 62 The FBI cannot read the messages, though, because they were encrypted. 63 Additionally, government officials have stated the Paris attackers likely used encrypted messaging services to plan the attacks. 64 Following the Paris attacks in November 2015, CIA Director John Brennan warned, “[t]here are a lot of technological capabilities that are available right now that make it exceptionally difficult both technically as well as legally for intelligence and security services to have the insight they need to uncover it.” 65 Terrorist recruiting, especially by the Islamic State of Iraq and al-Sham (ISIS), is increasingly occurring over encrypted *76 messaging services, too. 66 Director Comey has stated that ISIS operators in Syria are “recruiting and tasking dozens of troubled Americans to kill people, [using] a process that increasingly takes part through mobile messaging apps that are end-to-end encrypted, communications that may not be intercepted, despite judicial orders under the Fourth Amendment.” 67 Thus, encryption makes the government's counterterrorism mission much more difficult.

1. How Extensive is the Law Enforcement Risk?

In 2015, a total of 4148 wiretaps were authorized by federal and state judges. 68 Portable devices were the most frequently noted location in wiretap applications, as mobile communications, such as text messaging and application services, have become more prevalent. 69 In fact, 96% of all authorized wiretaps were designated as portable devices in 2015. 70 These are the devices that are increasingly becoming encrypted by default and several services are now using end-to-end encryption. Thus, as encryption becomes more widespread, law enforcement agencies will likely lose a robust source of evidence in criminal investigations and prosecutions.

Law enforcement has not reported a large number of encounters with encryption in recent years, but this is beginning to change. In 2014, just 26 wiretaps at the federal and state level were reported as being encrypted, of which only 5 could not
be decrypted. In 2015, just 13 wiretaps at the federal and state level were reported as being encrypted, of which 11 could not be decrypted. These statistics do not tell the full story, though, and understate the problem substantially. These data do not include intercepts that are authorized by the Foreign Intelligence Surveillance Act of 1978 (FISA), which authorizes surveillance of foreign intelligence and international terrorism. Also, law enforcement agencies will not use valuable time and resources trying to obtain a wiretap order when they know the communication will be encrypted and therefore the wiretap will not be productive, which is increasingly the case for connections between computers and websites, and for electronic messaging services. Thus, the statistics likely only represent the number of times that law enforcement has encountered encryption that thwarts an investigation when it is a surprise to law enforcement, because law enforcement will choose another method of investigation if it knows it will encounter encrypted communications when it obtains a wiretap order.

Cyrus Vance Jr., the Manhattan District Attorney, has testified that 74 cases in Manhattan from July 2014 to July 2015 were hindered because law enforcement was unable to access information on a device because of endpoint encryption. Vance's office later updated this number to 423 devices lawfully seized from October 2014 to October 2016 in which law enforcement's investigation was hampered by endpoint encryption. These data are just from one district in a single state. The Harris County District Attorney's Office, in which Houston, Texas is located, encountered between 8-10 encrypted devices every month in its criminal investigations in 2016. The Suffolk County District Attorney's Office, in which Boston, Massachusetts is located, encountered 151 encrypted devices in its criminal investigations during 2016. Law enforcement officials in Los Angeles, California were unable to access over 300 encrypted devices during criminal investigations in 2016. And, Wisconsin's Department of Justice encountered 68 encrypted devices in criminal investigations in 2016.

Further, in August 2016, Director Comey stated the FBI was unable to access 650 out of 5000 electronic devices that investigators attempted to search between October 2015 and August 2016. In March 2017, Director Comey updated these statistics and reported that the FBI was unable to access 1200 out of 2800 electronic devices that local police and federal agents sent to the FBI to search for evidence between October 2016 and December 2016. This indicates that law enforcement may encounter encryption quite frequently and that the problem may be large.

2. Is the Problem Overstated?

A recent report from Harvard University's Berkman Center for Internet and Society, “Don't Panic,” argues that the “going dark” problem may be overstated. Encryption, especially end-to-end encryption, may not become as widespread as law enforcement and intelligence officials fear. Furthermore, some believe that other means of surveillance will replace the information that is lost because of encryption, and that we are actually currently living in the “golden age for surveillance.”

End-to-end encryption may be unlikely to be widely adopted by industry because it conflicts with many companies' business models. Many companies rely on advertising revenue to subsidize the free content and services produced. Advertising is very dependent on user data for the production of targeted advertisements based on the user's Internet behavior and searches. End-to-end encryption would make user information unidentifiable, which would conflict with this business model. Thus, companies would likely lose revenue if they employed end-to-end encryption. Furthermore, cloud computing moves data and software and depends on ubiquitous connectivity, which allows the data
PREVENTING “GOING DARK”: A SOBER ANALYSIS AND..., 28 U. Fla. J.L. & Pub....

and software to be accessed through multiple platforms. This could not be done if end-to-end encryption was employed because the companies need access to the plaintext data. Therefore, end-to-end encryption may not be ubiquitous in the near future, but endpoint encryption does not face such obstacles and will likely become more widely adopted. However, the Office of the Director of National Intelligence (ODNI) has disputed that end-to-end encryption is unlikely to pose a significant problem. ODNI believes that there is already a trend developing of companies implementing end-to-end encryption and adversaries have already begun using these tools to avoid surveillance. Even if end-to-end encryption does not become widespread, it may still pose a significant threat if enough sophisticated malicious actors utilize products that adopt end-to-end encryption.

The Internet of Things (IoT) will greatly alter surveillance and present new opportunities to intercept information. IoT is the network of appliances and products, such as televisions, refrigerators, light bulbs, cameras, door locks, cars, watches, and other wearables that are connected and transmit data. As IoT becomes more widespread, as it is expected to do, this will enable the government to gain new avenues to have access to communications and data, especially metadata. ODNI strongly disputes that IoT and metadata could fill the gap of collection left by widespread encryption. Although metadata is unencrypted and will likely remain unencrypted--and provides a great deal of surveillance data already--it does not provide the content of communications. Metadata does not benefit law enforcement and intelligence agencies that are striving to prevent terrorist attacks and investigate and prosecute crimes in the same manner as actual content because the substance of communications remain unknown. This diminished usefulness is further exacerbated when expediency is important, which is often the case with counterterrorism efforts. Furthermore, many state and local law enforcement agencies will not have the resources and technology to conduct surveillance through IoT, and there will be many privacy implications from IoT that will prohibit law enforcement and intelligence agencies from lawfully collecting much of this data.

B. Private Sector and Security Concerns

The private sector and security officials have expressed numerous reasons why they oppose a lawful access requirement. The private sector and some cryptographers fear that the technological architecture that would guarantee law enforcement and intelligence agencies access would compromise user security and privacy. Building in lawful access would increase systems’ complexities, which would increase vulnerabilities because the new feature could interact with existing features in unintended and unknown ways. Also, the keys that would need to be retained by the companies, government, or third party would become targets for illicit actors to attack. For example, between 2004 and 2005, an unknown party exploited the intercept features of Greece's largest telephone service provider, which were designed to provide lawful access, and listened in on the calls of top Greek government officials before finally being detected. Thus, user security could be put at greater risk with a lawful access requirement. This is also very worrisome for the U.S. government because the United States is heavily dependent on cyber infrastructure, which makes it vulnerable to cyber-threats.

Additionally, surveillance by governments that have less robust legal processes as the United States would be made easier by the new technological architecture because U.S. products are used around the world. This would conflict with America's foreign policy interests at times when strong encryption would be favored because dissidents could use it to challenge authoritarian regimes. During the Arab Spring, for example, the State Department provided dissidents with encryption tools to help them shield their identities from repressive governments that were seeking to uncover those
participating in organized protests.  

This type of effort would be more difficult to accomplish if the United States mandated lawful access. Further, because U.S. products are used around the world, mandating lawful access would allow autocratic regimes to infringe on their citizens' privacy rights and enable these regimes to crackdown on dissidents. The United States would not have as much leverage in condemning such actions by repressive regimes if it demanded lawful access too because other countries would argue that they were legitimately pursuing law enforcement and intelligence goals through their actions. The United States would be able to argue that other countries' activities and surveillance laws are over-broad and repressive, though. In addition, authoritarian regimes would likely be able to bypass endpoint encryption regardless of whether lawful access was required if they have detained the user of the device because these regimes may resort to torture to obtain the keys to the device to find the desired information.

Also, the United States mandated lawful access would not be globally pervasive. There are 546 encrypted products from outside the United States. Thus, sophisticated illicit actors would be able to encrypt their communications regardless of whether the United States mandated lawful access. However, most illicit actors are not sophisticated. Many criminals end up getting caught because of flawed plans or carelessness. Therefore, the shift to overseas encryption products would likely not be widespread among illicit actors. In addition, U.S. companies dominate the market for technologies in which encryption capabilities are available. It is unknown, and perhaps doubtful, whether foreign companies may be able to significantly expand their market share if U.S. companies are forced to comply with a lawful access requirement.

The greatest potential harm from requiring lawful access likely stems from the possible decrease in the market share and economic viability of U.S. companies. U.S. intelligence has had a tremendous advantage in gathering information because a great deal of global communications transit in the United States. Following Edward Snowden's unauthorized disclosures of intelligence activities, foreign consumers have become concerned about U.S. surveillance. Foreign consumers may not want to use American products or online services if they believe their communications would be accessible to U.S. law enforcement or intelligence agencies. This could decrease U.S. companies' market share, which would mean less information would be transiting U.S. networks. Thus, U.S. intelligence agencies would have a more difficult time obtaining information. Also, a decrease in U.S. companies' market share would hurt their economic viability. U.S. technology companies have already lost between $35 and $180 billion in revenue over the 3-year period following the Snowden disclosures.

Additionally, U.S. companies face a tremendous threat from the theft of intellectual property through cyber espionage. General Keith Alexander, the former Commander of U.S. Cyber Command and Director of NSA, has stated that cyber espionage has resulted in the “greatest transfer of wealth in history.” U.S. companies lose about $250 billion per year because of intellectual property theft, global cybercrime costs companies about $114 billion per year worldwide, and $388 billion is lost globally when the costs of down time are taken into account. The threat of intellectual property theft and cybercrime could be exacerbated by a lawful access requirement that makes systems more vulnerable.

The decreased economic viability that could result from a lawful access requirement would diminish the economic strength of the United States, which is an important aspect of the role of the United States in the world. In 2014, Internet-related companies in the United States generated $966.2 billion in revenue, which accounted for 6% of real Gross Domestic Product. Economic strength enables countries to have political and military power and to have strong geopolitical influence. Therefore, it is important to consider the economic interests of U.S. businesses when considering whether a lawful access requirement is appropriate.
Many who oppose a lawful access requirement still recognize that law enforcement and intelligence agencies would be harmed. Instead of a lawful access requirement, they propose that law enforcement and intelligence agencies be granted greater authorities, resources, and capabilities to obtain the necessary information to keep the country secure through alternative means. A prominent alternative to a lawful access requirement would be for law enforcement and intelligence agencies to focus on legally exploiting vulnerabilities in targeted devices to conduct surveillance.  

1. Is Legally Exploiting Vulnerabilities a Good Alternative to a Lawful Access Requirement?

Legally exploiting vulnerabilities could help law enforcement and intelligence agencies obtain the information they need without mandating a lawful access requirement. Those who advocate for this approach believe that “the technique is preferable for conducting wiretaps against targets when compared to other possible methods of wiretapping,” such as a lawful access requirement.  

Another variation of this approach is that law enforcement and intelligence agencies should have the ability to require companies to install software on consumers' devices that “will copy and forward all communications that are sent or received through *84 that device” to government authorities. However, exploiting vulnerabilities raises serious legal, business, transparency, and scalability issues. If the government must seek the technical assistance of a company to push software to a device to facilitate surveillance, this must adhere to the technical assistance provisions of either FISA—for cases involving foreign powers and agents of foreign powers—or the Wiretap Act—for criminal cases. FISA states that the Foreign Intelligence Surveillance Court (FISC) shall direct that,  

upon the request of the applicant, a specified communication or other common carrier, landlord, custodian, or other specified person ... furnish the applicant forthwith all information, facilities, or technical assistance necessary to accomplish the electronic surveillance in such a manner as will protect its secrecy and produce a minimum of interference with the services that such carrier, landlord, custodian, or other person is providing that target of electronic surveillance.  

Similarly, the Wiretap Act requires that upon a court order,  

a provider of wire or electronic communication service, landlord, custodian or other person shall furnish the applicant forthwith all information, facilities, and technical assistance necessary to accomplish the interception unobtrusively and with a minimum of interference with the services that such service provider, landlord, custodian, or person is according the person whose communications are to be intercepted.  

Although both of these statutes are very broad, and the FISA provision is largely uninterpreted, the Supreme Court's interpretation of the Wiretap Act's technical assistance provision presents a potential difficulty for the government to be able to compel a company to push *85 software to a device to facilitate surveillance. In determining that a company had to assist law enforcement in the installation of a pen register device, Supreme Court in New York Telephone Co. reasoned that,  

it can hardly be contended that the Company, a highly regulated public utility with a duty to serve the public, had a substantial interest in not providing assistance. Certainly the use of pen registers is by no
means offensive to it. The Company concedes that it regularly employs such devices without court order for the purposes of checking billing operations, detecting fraud, and preventing violations of law. It also agreed to supply the FBI with all the information required to install its own pen registers. Nor was the District Court's order in any way burdensome. The order provided that the Company be fully reimbursed at prevailing rates, and compliance with it required minimal effort on the part of the Company and no disruption to its operations. 128

This indicates that courts will consider whether a company is a public utility or private entity, whether the company normally performs the activity the government is asking in its course of business, and whether the request is unduly burdensome for the company in determining whether a company must push software to a device to facilitate surveillance. Courts could consider the fact that companies regularly push software updates to consumers and that the government will compensate companies financially 129 for the assistance as indicative that companies must be forced to provide assistance to the government in pushing software to a device to facilitate surveillance. However, courts may determine that private companies that do not have a public service obligation and that do not engage in pushing software updates to consumers that intentionally create vulnerabilities would be unduly burdened by degrading the companies' trust relationship with its consumers.

The Eastern District of New York in In re Apple, Inc. came to the latter conclusion in ruling that Apple could not be compelled to bypass its endpoint encryption on a defendant's device by the All Writs Act, 130 which has come to be understood as authorizing a federal court to issue a writ directing a third-party to the underlying litigation to provide reasonable technical assistance to the government to facilitate the execution of a valid search warrant. 131 The court reasoned that Apple, a private company (not a public utility with a duty to serve the public), had an interest in not complying with the government because of its reliance on its reputation as a company that safeguards user data. 132 Further, the court determined that

unlike the pen register at issue in N.Y. Tel. Co., the assistance the government [sought] here--bypassing a security measure that Apple affirmatively markets to its customers--is not something that Apple would normally do in the conduct of its own business and is, at least now, plainly “offensive to it.” 133

Finally, the court determined that bypassing the endpoint encryption diverted man-hours, software, and hardware from Apple's normal business operation and therefore constituted an unreasonable burden. 134

The opinion in In re Apple, Inc. indicates that when faced with the question of whether a company should be compelled under FISA or the Wiretap Act to provide assistance to the government in pushing software to a device to facilitate surveillance, a court may similarly determine that such assistance would constitute an undue burden. Even if courts determine that companies must provide assistance to the government in pushing software to a device to facilitate surveillance, this could seriously degrade the trust relationship consumers have with the companies. This could lead consumers to stop updating their software, which in turn would degrade security because companies typically push software updates to consumers in part to improve the security of systems and applications and fix vulnerabilities. Thus, consumers could actually be left more vulnerable to cyber-threats from malicious actors if they chose to stop receiving software updates out of fear--whether legitimate or not--that the updates may be facilitating surveillance.
Also, transparency with the public will be diminished if the government is forced to rely on legally exploiting vulnerabilities. Currently, companies provide transparency reports regarding government requests for data, including FISA orders after the passage of the USA FREEDOM Act permitted such disclosures. The government also provides reports on FISA orders and court orders for communications interception. These levels of transparency are unlikely to occur with the government's exploitation of vulnerabilities because companies would likely be reticent to disclose the number of vulnerabilities being exploited in their products, which they would rather like to tout as being secure.

Further, some advocates of this approach argue that the government should disclose all, or nearly all, vulnerabilities discovered to companies so that these vulnerabilities may be fixed. While this would help companies to further improve the security of their systems and applications, law enforcement and intelligence agencies would in turn be unable to exploit those same vulnerabilities to produce information in the future. Certainly, the government's role in vulnerability disclosure is already a hotly debated issue, and forcing law enforcement and intelligence agencies to rely more heavily on exploiting vulnerabilities will only add to the complexity of this difficult issue.

Finally, relying on legally exploiting vulnerabilities is not yet a practical approach on a large-scale. This approach would likely not be able to provide the same level of data available for collection by law enforcement and intelligence agencies as is currently available. In addition, it is very doubtful that state and local authorities have the technical sophistication and resources to develop the necessary vulnerabilities or to compensate companies for their assistance.

While legally exploiting vulnerabilities does present opportunities for law enforcement and intelligence agencies, the approach presents numerous difficulties as a replacement for a lawful access requirement, and likely requires Congress to pass comprehensive legislation to permit and regulate its use. Thus, legally exploiting vulnerabilities does not present a sufficient alternative to a lawful access requirement.

2. Would Security and Privacy Really Be Compromised by a Lawful Access Requirement?

Benjamin Wittes, a senior fellow at the Brookings Institution and cofounder of Lawfare, has raised the point that major Internet companies currently have the ability to decrypt information. Google has the ability to decrypt Gmail and Gchat communications because this allows Google to target users for advertisements. Also, Gmail is able to filter spam, which can contain malware, because Google can read emails' plaintext, which would not be possible with end-to-end encryption. Further, Google offers the full text search of files stored in the cloud, which requires access to plaintext, too, and could not occur with end-to-end encryption. There have not been security issues with Google's services thus far. Despite a lack of security problems, Google has announced that it will offer a messaging application, Allo, that will use opt-in (not default) end-to-end encryption, and video calling application, Duo, that will use default end-to-end encryption. Facebook's Messenger used to be able to be decrypted by Facebook. Although the company has started testing opt-in (not default) end-to-end encryption for Messenger, Facebook has not indicated this move is because the service was insecure. In addition, Apple continues to enable iCloud backups to occur in a manner that can be decrypted. Once again, there have not been security issues with the service. Thus, there does not seem be evidence that these major Internet companies' services are insecure because they have the ability to decrypt information. Some of the fears of insecurity seem to actually stem from a disdain for lawful surveillance. However, lawful surveillance is not a cybersecurity risk.
Lawful surveillance is only conducted by the United States following rigorous legal scrutiny. The government conducts surveillance of the *90 content of communications pursuant to Title III and FISA, and obtains stored communications pursuant to the Stored Communications Act (SCA), 151 which is part of the Electronic Communications Privacy Act (ECPA). 152 Under Title III, in order to obtain a court order to intercept wire, oral, and electronic communications, the government is required to establish:

(a) there is probable cause for belief that an individual is committing, has committed, or is about to commit [an enumerated crime]; (b) there is probable cause that particular communications concerning that offense will be obtained through such interception; (c) normal investigative procedures have been tried and have failed or reasonably appear to be unlikely to succeed if tried or to be too dangerous; [and] (d) ... there is probable cause for belief that the facilities from which, or the place where, the wire, oral, or electronic communications are to be intercepted are being used, or are about to be used, in connection with the commission of such offense, or are leased to, listed in the name of, or commonly used by such person. 153

Under FISA, the government is required to establish probable cause that the “target of the electronic surveillance is a foreign power or an agent of a foreign power”; probable cause that “each of the facilities or places at which the electronic surveillance is directed is being used, or is about to be used, by a foreign power or an agent of a foreign power”; and that the proposed minimization procedures 154 meet the statutory requirements. 155 The government may compel the disclosure of the contents of stored wire or electronic communication “that is in electronic storage in an electronic communications system for one hundred and eighty days or less” by an electronic communication service only upon receipt of a warrant under the SCA. 156 Also, under the SCA, the government may compel a provider of remote computing service to disclose the contents of stored wire or electronic communication with a warrant, or by obtaining an administrative subpoena or court order and *91 providing prior notice to the subscriber or customer. 157 To obtain a court order under the SCA, the government must “offer[] specific and articulable facts showing that there are reasonable grounds to believe that the contents of a wire or electronic communication, or the records or other information sought, are relevant and material to an ongoing criminal investigation.” 158 These statutory requirements ensure that people are provided robust privacy protections whenever their Fourth Amendment rights may be implicated. The concerns that user security and privacy would be compromised and that intellectual property theft would be exacerbated may be overblown.

3. Is There Really Consumer Pressure Driving Companies Toward More Encryption?

A lawful access requirement's potential harm to the market share and economic viability of U.S. companies is the strongest argument for not mandating such a requirement. Certainly, the Snowden disclosures had a negative impact on U.S. companies in their immediate aftermath, but confidence in U.S. companies seems to have rebounded. In the fourth quarter of 2016, Facebook reported that it had 1.86 billion users who logged on at least once a month. 159 This is up from the 1.16 billion users who logged on at least once a month in the second quarter of 2013, when Snowden first disclosed NSA programs. 160 In 2015, 70% of Facebook users logged on daily, which is an increase from the 63% who logged on daily in August 2013, in the immediate aftermath of the Snowden disclosures. 161 Also, Google announced in February 2016 that Gmail has over one billion active users. 162 This is an increase from the 900 million active users Gmail had in 2015, when it became the most popular email service in the United States, and from the 425 million active users Gmail 92 had in 2012, prior to the Snowden disclosures. 163 Of these Gmail users, 75% accessed their accounts
on mobile devices, too. These statistics do not indicate that consumers are deeply distrustful of U.S. companies and refuse to use U.S. products and services.

Further, a recent study surveying 1510 participants, including both information technology security experts and non-experts, from the United States, United Kingdom, and Germany, found that privacy and security only play a minor role in people's decisions to use a particular mobile instant messenger. The primary reason that participants gave for using a mobile instant messenger was whether friends were using the messenger. 46.1% of participants from the United States, 48.2% of participants from the United Kingdom, and 54.9% of participants from Germany stated this was the main reason they used a particular mobile instant messenger. On the other hand, only a small percentage of participants stated the main reason they used a mobile instant messenger was because of privacy and security. Only 5.6% of participants from the United States, 3.4% of participants from the United Kingdom, and 13.1% of participants from Germany stated this was the main reason that they used a particular mobile instant messenger.

Consumer pressure may not truly be driving the move toward more encryption; instead, companies may be pressuring each other to move in this direction. Consumers seem to care more about being able to be connected to friends, having easy to use and reliable products, and having sleek interfaces and useful applications, and seem willing to sacrifice some privacy and security in exchange. If consumer pressure is not driving this shift, then perhaps the fear that U.S. companies will lose market share and that the economic viability of U.S. companies would be hurt by a lawful access requirement is overstated. As long as U.S. companies continue to lead in the areas consumers care about the most, while still making the most secure products and services possible that comply with a lawful access requirement, U.S. companies will likely continue to dominate the technology market.

*93 C. The U.S. Should Mandate a Lawful Access Requirement

There should be a lawful access requirement. Encryption presents great harms to law enforcement and intelligence agencies. Without a lawful access requirement, lawful surveillance would become extremely difficult and cyberspace would become a large ungoverned space. This would drastically limit the ability to intercept terrorist communications, which would inhibit the ability to prevent attacks, and would hinder law enforcement from performing its mission of investigating and prosecuting crimes. Ultimately, crimes would go unsolved, criminals would not be brought to justice, and the United States and its allies would be made more vulnerable to terrorist attacks.

While some cybersecurity concerns and vulnerabilities may arise from creating the architecture to provide lawful access, the fact that companies currently have the ability to decrypt information demonstrates that this can in fact be done in a secure manner. Authoritarian regimes already invade their citizens' privacy and seek to crack down on dissidents, and will implement access requirements of their own regardless of what the United States chooses to do. The United States can still be critical of other nations' human rights abuses and less robust legal processes even if the United States implements a lawful access requirement. Furthermore, while some criminals and terrorists will find ways to use encrypted communication no matter what, many criminals and terrorists will not be sophisticated enough to do this. The United States could use other resources and focus on legally exploiting vulnerabilities in targeted devices to conduct surveillance on these more sophisticated actors, which is more feasible than trying to focus on legally exploiting vulnerabilities for all criminal and terrorist targets. Finally, U.S. companies have rebounded since the Snowden disclosures, and do not appear to face significant market share and economic vulnerabilities from a lawful access requirement. Thus, the United States should mandate a lawful access requirement, and it is technologically feasible to securely achieve this.
V. HOW A LAWFUL ACCESS REQUIREMENT COULD BE ACHIEVED

There are numerous techniques through which a lawful access requirement could be achieved. Technological innovations could be *developed, users could be given greater control, or industry incentives could be altered through conditional liability protections.

A. Possible Technological Innovations

1. Weaken Encryption Systems or Renew the Push for a Key Escrow System

While weakening encryption systems would create vulnerabilities that could be exploited by illicit actors in addition to law enforcement and intelligence agencies, there are several technological means that could be used in a secure manner. The government could once again seek to institute a key escrow system, in which the government or trusted third party would hold the keys. A key escrow system would not weaken encryption and the keys would be made accessible only upon lawful process, which would assuage civil liberties concerns. However, a key escrow system would increase the surface area of the underlying system, which would increase the system's susceptibility to attacks.

2. Mandatory Biometric Encryption

Paul Rosenzweig, a Senior Advisor to The Chertoff Group and former Deputy Assistant Secretary for Policy at the Department of Homeland Security, has pointed out that mandatory biometric encryption would allow for strong endpoint encryption while providing lawful access. This strategy would require that manufacturers use encryption with a biometric lock, such as a fingerprint, rather than a password. Biometric encryption provides more robust security for users than passwords, but is likely not protected by the Fifth Amendment, unlike passwords.

The Fifth Amendment privilege protects an individual from being “compelled in any criminal case to be a witness against himself.” The Supreme Court has interpreted this as protecting an individual from being compelled to give testimony that is self-incriminating. Cases typically focus on whether a statement is testimonial, as compulsion and the incriminating nature of documents are rarely in doubt. The Supreme Court has held that it is testimonial implicating the Fifth Amendment when the government compels an individual to use the contents of his own mind to communicate a fact or disclose information. In *In re Grand Jury Subpoena Duces Tecum Dated March 25, 2011*, the government sought to compel the defendant to decrypt and produce the contents of hard drives that the government suspected contained child pornography. However, the Eleventh Circuit Court of Appeals ruled that compelled password decryption violates the Fifth Amendment because it involves requiring the defendant to use the defendant's mind to incriminate himself. On the other hand, the Supreme Court in *Schmerber* ruled that only compelled communications or testimony are prohibited by the privilege against self-incrimination, and that physical evidence, such as fingerprints, from a suspect is not subject to Fifth Amendment protection.

Therefore, biometric encryption would both enhance user security and enable law enforcement to obtain evidence in criminal investigations when law enforcement has detained a suspect. However, this solution would not address end-to-end encryption, which also presents a difficult challenge, especially in counterterrorism efforts where end-to-end encryption is being used by ISIS in recruiting and has likely been used to plan attacks. End-to-end encryption may not become ubiquitous, though, because it conflicts with many companies' business models.
3. Split Key Encryption

Professor Geoffrey Corn's proposal of split key encryption is a more promising potential solution. Encryption keys can be split into separate parts and stored discretely. Split keys can be further secured by being encapsulated inside other encrypted data. The split keys placed inside other encrypted data can be placed in the possession of multiple entities, who would need to cooperate with each other to unlock the set of keys to ultimately decrypt the sought after information. Thus, all parties involved would need to be satisfied by the legal order to facilitate surveillance, and the fact that the key would be split with the discrete parts stored with separate entities would make it very difficult for illicit actors to obtain the keys through cyber-theft. This technique would increase complexity, though, which could create other vulnerabilities that could be exploited. This system would require further development to reduce complexity and to determine how keys would be obtained when communication occurs across different countries.

4. Cryptographic Envelopes

Also, Matt Tait's, the CEO and founder of Capital Alpha Security and former Government Communications Headquarters (GCHQ) information security specialist, suggestion of cryptographic envelopes presents a very appealing technological solution. In this approach, the key to the device's drive would be placed inside a cryptographic envelope. The drive would then be able to be decrypted either by the user's password or by opening the cryptographic envelope. The cryptographic envelope would be addressed to an entity using the entity's public key, which is published and is what encrypts the information, and would be sealed using strong cryptography. The cryptographic envelope could only be opened using that entity's corresponding private key, which is known only to that entity. For additional protection, the cryptographic envelope could be placed inside other cryptographic envelopes. Thus, a cryptographic envelope could be addressed to the FBI using the FBI's public key, and then placed inside another cryptographic envelope that could be addressed to the manufacturer using the manufacturer's public key. In this scenario, neither the FBI nor the manufacturer could unilaterally decrypt the information because the FBI could not open the outer cryptographic envelope because it does not have the manufacturer's private key, and the manufacturer could not open the inner cryptographic envelope because it does not have the FBI's private key. Decryption of the information would instead require the FBI and manufacturer to work together following applicable legal process, and other cryptographic envelopes could even be layered on top for further protection. Therefore, this system would enable law enforcement and intelligence agencies to obtain information upon legal order, and would “block access to thieves, foreign governments, hackers, criminal misuse by rogue police officers without a warrant, or misuse by staff working for the device's manufacturer.”

B. Allowing the User to Choose

Ariel Rabkin, a visiting fellow at the American Enterprise Institute, has raised the possibility that users could be empowered to control whether they use encryption. Manufacturers could be mandated to produce devices that will provide lawful access, but users could be able to alter devices to add encryption that would not enable lawful access. This would remove the increasing tendency for devices to be encrypted by default. Of course, criminals and terrorists will likely alter devices to use encryption under this strategy, but many illicit actors will not be sophisticated enough to do so, and the United States could concentrate other resources on combating these sophisticated malicious actors.
An alternative approach, put forward by Benjamin Wittes, would be for encrypted services to be mandated to offer an “Emergency Access Mode.” The “Emergency Access Mode” would enable the company to unlock the device or decrypt communications upon the death or incapacity of the consumer, or upon applicable legal process. Consumers would be prompted to determine whether they wanted to use the “Emergency Access Mode” or use endpoint or end-to-end encryption when they first use the device or service. Under this strategy, government officials, such as Director Comey, would educate people about the benefits of choosing the “Emergency Access Mode” by raising awareness of cases like the unsolved murder that occurred in Louisiana. This approach would also remove the increasing tendency for devices to encrypt data by default, and many people may choose to use the “Emergency Access Mode” when they are made aware of the risks of encryption. Companies could be required to disclose which consumers elect to not use the “Emergency Access Mode,” and instead elect to use encryption that cannot be deciphered, even upon a lawful order. This information could be cross-referenced with other data, such as the no-fly list and sex offender registry, which could be useful for law enforcement and intelligence agencies. However, this requirement would likely face stiff opposition from privacy advocates, and would therefore be unlikely to be passed by Congress. Ultimately, these approaches are dependent on users, who in the end may choose to use encrypted services, which would render these approaches unhelpful to law enforcement and intelligence agencies.

C. Shifting Industry Incentives

Finally, Benjamin Wittes also put forward the idea that Congress could condition civil liability immunity “on whether companies maintain the technical capacity to deliver interpretable signal in response to lawful wiretap orders.” Currently, section 230(c)(1) of the Communications Decency Act (CDA) grants immunity to online service providers and users from civil suits for third-party conduct, and is very important to Internet companies. Congress could condition this immunity on companies' ability to respond to lawful wiretap orders. Companies would still be able to pursue encryption that cannot be deciphered upon lawful process, but the potential loss of liability protection may incentivize companies to maintain the ability to comply with wiretap orders. This approach would open these companies up to frivolous lawsuits, though, which could impair business in unintended ways. Also, companies may be able to rely on other legal defenses, which would allow them not to alter their behavior in the manner this approach intends.

VI. THE PATH FORWARD

Ultimately, there should be a lawful access requirement. Congress should legislate that companies be required to have the capacity to deliver decrypted information upon applicable legal process, and not enable users to alter devices. This legislative requirement should enable the technology industry to innovate to achieve the mandate, which would achieve a better outcome than allowing the user to choose or trying to shift industry incentives through conditional liability protection. Cryptographic envelopes, split key encryption, and possibly even biometric encryption, depending on how widespread end-to-end encryption is adopted, all present potential solutions for achieving this technological mandate in a secure manner. U.S. companies will likely compete to provide the most secure products and services for consumers that comply with this mandate, and will likely develop additional technological means to comply with the requirement in a secure manner. These companies are unlikely to lose market power or economic viability because of a lawful access requirement. This solution would enable law enforcement and intelligence agencies to continue pursuing their security missions in an effective manner, and would maintain cybersecurity.
Footnotes

a1 Brumley Next Generation Senior Graduate Fellow, Intelligence Studies Project, Robert S. Strauss Center for International Security and Law. B.A. 2013, Rice University; J.D. and Master of Public Affairs Candidate 2018, The University of Texas School of Law and Lyndon B. Johnson School of Public Affairs.


2 CHERTOFF GROUP, supra note 1, at 3.


4 Id.

5 Id.


7 Id.

8 See generally id.

9 See 22 C.F.R. § 121, Category XIII(b) (1995) (listing the cryptographic items on the munitions list); id. § 125.1 (requiring a license for the export of items listed on the munitions list).

10 See, e.g., John Markoff, Paper on Codes Is Sent Despite U.S. Objections, N.Y. TIMES (Aug. 9, 1989), http://www.nytimes.com/1989/08/09/us/paper-on-codes-is-sent-despite-usobjections.html (“Over the past 12 years the National Security Agency has consistently opposed the publication or transmission of research on encryption technology. The agency is concerned that advances in cryptography will make it harder to break coded transmissions sent by foreign intelligence agents in the United States to their governments.”).


12 See Statement by The Press Secretary, WHITE HOUSE 1 (Apr. 16, 1993), https://www.epic.org/crypto/clipper/white_house_statement_4_93.html (acknowledging law enforcement’s concerns about losing the ability to lawfully intercept communications when announcing the Clipper Chip proposal).

13 See Peter Swire & Kenesa Ahmad, ‘Going Dark’ Versus a ‘Golden Age for Surveillance,’ CTR. FOR DEMOCRACY & TECH. (Nov. 28, 2011), https://cdt.org/blog/%E2%80%98going-dark%E2%80%99-versus-a-%E2%80%98golden-age-for-surveillance%E2%80%99/ (discussing the origins of law enforcement’s and intelligence agencies’ worries about losing surveillance capabilities before arguing that law enforcement and intelligence agencies have a greater access to data today than ever before).

14 Id.


16 CASTRO & MCQUINN, supra note 11, at 7-8.
PREVENTING “GOING DARK”: A SOBER ANALYSIS AND...

17  Id. at 7-8.

18  Id.


20  Id.

21  See id. (noting that while the Clipper Chip was voluntary, the government hoped the marketplace would adopt the system).


27  Id. at 3.


29  Id. § 1002(b)(3).


31  Id.

32  See 22 C.F.R. § 121, Category XIII(b) (1995) (listing the cryptographic items on the munitions list); id. § 125.1 (requiring a license for the export of items listed on the munitions list).


34  Bernstein v. U.S. Dep't of Justice, 176 F.3d 1132, 1145-46 (9th Cir. 1999), reh'g granted, opinion withdrawn, 192 F.3d 1308 (9th Cir. 1999). The court also expressed that “[g]overnment efforts to control encryption thus may well implicate not only the First Amendment rights of cryptographers intent on pushing the boundaries of their science, but also the constitutional rights of each of us as potential recipients of encryption's bounty.” Id. at 1146. The case did not continue because the government revised its policy.


36  OLSEN ET AL., supra note 3, at 4-5.

37  Swire & Ahmad, supra note 13.
38. CHERTOFF GROUP, supra note 1, at 1.

39. Id.

40. Id.


42. See id. (noting Google's switch to default encryption).

43. CHERTOFF GROUP, supra note 1, at 1.

44. Id.

45. Id.


47. See Charlie Savage, U.S. Tries to Make It Easier to Wiretap the Internet, N.Y. TIMES (Sept. 27, 2010) http://www.nytimes.com/2010/09/27/us/27wiretap.html (discussing the FBI's initial efforts to address the growing concerns that investigators will lose the ability to intercept communications they are lawfully authorized to intercept).


49. See generally FEDERAL BUREAU OF INVESTIGATION SITUATIONAL INFORMATION REPORT, GOING DARK: LAW ENFORCEMENT PROBLEMS IN LAWFUL SURVEILLANCE (2011) (explaining the problems that new technologies are posing for lawfully-authorized electronic surveillance).

50. See Anna Mulrine, New Encryption Technology is Aiding Terrorists, Intelligence Director Says, CHRISTIAN SCI. MONITOR (Apr. 25, 2016), http://www.csmonitor.com/USA/Politics/monitor_breakfast/2016/0425/New-encryption-technology-is-aiding-terrorists-intelligence-director-says (reporting on James Clapper's, the Director of National Intelligence, comments regarding encryption inhibiting the Intelligence Community's ability to collect intelligence, especially against terrorists); CIA Director John Brennan on 60 Minutes, 60 MINUTES (Feb. 14, 2016), http://www.cbsnews.com/news/cia-director-john-brennan-60-minutes-scott-pelley/ (stating that encryption has hampered the CIA's ability to collect intelligence on ISIS); Michael Isikoff, NSA Chief: ‘Paris Would not have Happened’ Without Encrypted Apps, YAHOO (Feb. 17, 2016), https://www.yahoo.com/news/nsa-chief-paris-would-not-have-happened-without184040933.html (Admiral Michael Rogers, Commander of U.S. Cyber Command and Director of NSA, warned encryption is making it more difficult for the NSA to intercept communications and fulfill its mission).

51. Adam Segal & Alex Grigsby, How to break the deadlock over data encryption, WASH. POST (Mar. 13, 2016), https://www.washingtonpost.com/opinions/how-to-break-the-deadlock-over-data-encryption/2016/03/13/e677fb78-d110-11e5-88cd-753e80cd29ad_story.html?utm_term=.0a2edc86e8 (“The challenge of ‘going dark’ affects state and local law enforcement the most: They are the least likely to have the resources and technical capabilities to decrypt data relevant to an investigation.”).

52. OLSEN ET AL., supra note 3, at 6.


Administrative Office of the U.S. Courts, Wiretap Report 2014, UNITED STATES COURTS (Dec. 31, 2014), http://www.uscourts.gov/statistics-reports/wiretap-report-2014; Administrative Office of the U.S. Courts, Wiretap Report 2015, supra note 68. Encryption was reported for the first time in 2015 in one federal wiretap that was conducted in 2014. Officials were unable to decrypt the communication in that intercept.


MANHATTAN DISTRICT ATTORNEY'S OFFICE, supra note 77, at 9-10.


See generally OLSEN ET AL., supra note 3.

Id. at 11.

Id. at 13-15 (discussing the surveillance opportunities that the Internet of Things may present); Swire & Ahman, supra note 23, at 463-73 (arguing that we are currently in the “golden age for surveillance”).

OLSEN ET AL., supra note 3, at 10.

Id. at 10-11.

See Create Gmail Ads with Templates, GOOGLE (last visited Aug. 1, 2016), https://support.google.com/adwords/answer/6105478 (discussing targeted Gmail advertisements); How to target Facebook Ads, FACEBOOK (last visited Aug. 1, 2016), https://www.facebook.com/business/a/online-sales/ad-targeting-details (boasting 89% accuracy for targeted advertising campaigns based on Facebook's information on users' location, demographics, interests, and behaviors).

OLSEN ET AL., supra note 3, at 11.
PREVENTING “GOING DARK”: A SOBER ANALYSIS AND..., 28 U. Fla. J.L. & Pub....

93    Id.
94    See Letter from the Office of the Director of National Intelligence to Senator Ron Wyden (May 5, 2016), 1-2, https://www.wyden.senate.gov/download?id=3F716160-095E420E-93F3-849453EB61B2&download=1 (asserting that the increased prevalence of encryption has already hampered law enforcement and intelligence collection activities and that the problem is only growing).
95    Id.
97    Id. at 13.
98    Id. Metadata, such as e-mail addresses and telephone numbers, is the data that describes and gives information about other data.
100   Comey, supra note 53.
103   The government is not a monolithic actor in this debate, and several former national security and intelligence officials oppose a lawful access requirement.
104   See generally HAROLD ABELSON ET AL., KEYS UNDER DOORMATS: MANDATING INSECURITY BY REQUIRING GOVERNMENT ACCESS TO ALL DATA AND COMMUNICATIONS (2015) (arguing against a lawful access requirement because of cybersecurity concerns).
105   Id. at 15-17.
106   Id.
107   Vassilis Prevelakis & Diomidis Spinellis, The Athens Affair, IEEE SPECTRUM (June 29, 2007), http://spectrum.ieee.org/telecom/security/the-athens-affair (detailing the surreptitious wiretapping of Vodafone Greece that allowed the communications of prominent Greek politicians to be intercepted by hackers).
109   See Andrea Peterson, The NSA is Trying to Crack Tor. The State Department is Helping Pay for It, WASH. POST (Oct. 5, 2013), https://www.washingtonpost.com/news/the-switch/wp/2013/10/05/the-nsa-is-trying-to-crack-tor-the-state-department-is-helping-pay-for-it/ (reporting on the State Department's effort to teach activists and journalists to use Tor and other counter-surveillance technologies during the Arab Spring).
111   Id. at 6.
112   See Wang, supra note 108 (detailing the prominence of U.S. technologies around the world).
John Markoff, *Internet Traffic Begins to Bypass the U.S.*, N.Y. TIMES (Aug. 29, 2008), http://www.nytimes.com/2008/08/30/business/30pipes.html?pagewanted=print&r=0 (noting General Michael Hayden's, former Director of the CIA and Director of the NSA, testimony before the Senate Judiciary Committee stating that “[b]ecause of the nature of global telecommunications, we are playing with a tremendous home-field advantage, and we need to exploit that edge”).


See, e.g., id.


Id.


Id. at 64.

Herb Lin, *A Biometric Approach as a Partial Step Forward in the Encryption Debate*, LAWFARE (Dec. 3, 2015, 3:22 AM), https://www.lawfareblog.com/biometric-approach-partialstepforward-encryption-debate. Companies already push software updates to consumers and this variation of exploiting vulnerabilities would allow government authorities to create a vulnerability it could exploit on a specific device through this mechanism. *Id.*


These provisions do not require technical assistance in decrypting signal, which would solve the “going dark” problem, because of the interaction with CALEA's exception that “[a] telecommunications carrier shall not be responsible for decrypting, or ensuring the government's ability to decrypt, any communication encrypted by a subscriber or customer, unless the encryption was provided by the carrier and the carrier possesses the information necessary to decrypt the communication.” 47 U.S.C. § 1002(b)(3) (2012).

A pen register is a “device or process which captures the incoming electronic or other impulses which identify the originating number or other dialing, routing, addressing, and signaling information reasonable likely to identify the source of a wire or electronic communication, provided, however, that such information shall not include the contents of any communication.” 18 U.S.C. § 3127(4) (2012).
See 18 U.S.C. § 2518(4) (2012) (requiring the government to provide compensation for assistance in intercepting communications); 18 U.S.C. § 3124(c) (2012) (requiring the government to provide compensation for the facilitation of the installation of pen register and trap and trace devices); 18 U.S.C. § 3125(d) (2012) (requiring the government to provide compensation for the facilitation of emergency pen register and trap and trace device installation); 50 U.S.C. § 1802(a)(4) (2012) (requiring the government to provide compensation to carriers for furnishing aid in carrying out electronic surveillance authorized by the Attorney General without a court order); 50 U.S.C. § 1805 (c)(2)(D) (2012) (requiring the government to provide compensation for the provision of technical assistance in cases involving foreign powers and agents of foreign powers); 50 U.S.C. § 1824(c)(2)(D) (2012) (requiring the government to compensate a person for facilitating a physical search); 50 U.S.C. § 1842(d)(2)(B)(iii) (2012) (requiring the government to provide compensation for reasonable expenses incurred in assisting with the installation of pen register and trap and trace devices for foreign intelligence and international terrorism investigations); 50 U.S.C.A. § 1861(j) (West 2016) (requiring the government to compensate a person for reasonable expenses incurred for producing tangible things or providing information, facilities, or assistance in gaining access to certain business records for foreign intelligence and international terrorism investigations); 50 U.S.C. § 1881a(h)(2) (2012) (requiring the government to compensate electronic service providers for assisting the government to acquire communications related to foreign intelligence information); 50 U.S.C. § 1881b(c)(5)(d) (2012) (requiring the government to provide compensation at the prevailing rate to electronic service providers for assisting in the acquisition of communications and data of targeted U.S. persons outside the United States).

See Marshall Erwin, Lawful Hacking After the Encryption Debate, JUST SECURITY (Oct. 15, 2015, 1:00 PM), https://www.justsecurity.org/26849/lawful-hacking-encryption-debate/ (advocating that legally exploiting vulnerabilities is a better alternative to a lawful access requirement, but acknowledging the potential for a loss of transparency).


See Sara Sorcher & Malena Carollo, Influencers: FBI Should Disclose San Bernardino iPhone Security Hole to Apple, CHRISTIAN SCI. MONITOR (Mar. 24, 2016), http://passcode.csmonitor.com/influencers-applevsfbi (citing several experts advocating for vulnerabilities to be disclosed as quickly as possible).

See Michael Daniel, Heartbleed: Understanding When We Disclose Cyber Vulnerabilities, WHITE HOUSE: WHAT’S HAPPENING (Apr. 28, 2014, 3:00 PM), https://www.whitehouse.gov/blog/2014/04/28/heartbleed-understanding-when-we-disclose-cyber-vulnerabilities (“Disclosing a vulnerability can mean that we forego an opportunity to collect crucial intelligence that could thwart a terrorist attack[,] stop the theft of our nation’s intellectual property, or even discover more dangerous vulnerabilities that are being used by hackers or other adversaries to exploit our networks.”).

See Hearing on Going Dark: Lawful Electronic Surveillance in the Face of New Technologies, supra note 48, at 2 (statement of Rep. Tim Griffin, Member, H. Subcomm. on Crime, Terrorism, and Homeland Security) (acknowledging that state and local law enforcement agencies do not have the same financial resources or technical expertise as federal law enforcement, and encouraging others to take this into account in analyzing the “going dark” problem); Robin Hanson, Can Wiretaps Remain Cost Effective?, COMM. OF THE ACM, Dec. 1994, at 13, 13-15 (recognizing that police are price sensitive to the cost of surveillance activities and must factor cost into the decision whether to pursue that method of investigation).


Id.


OLSEN ET AL., supra note 3, at 11.

Wittes, Five Hard Encryption Questions, supra note 142.

Andy Greenberg, With Allo and Duo, Google Finally Encrypts Conversations End-to-End, WIRED (May 18, 2016, 3:23 PM), https://www.wired.com/2016/05/allo-duo-google-finally-encrypts-conversations-end-end/.

Wittes, Five Hard Encryption Questions, supra note 142.


Wittes, Five Hard Encryption Questions, supra note 142.


Minimization procedures are a set of rules that dictate how a government agency will limit the accessibility, retention, and dissemination of inadvertently acquired material concerning U.S. persons who are not the target of the surveillance. 50 U.S.C. § 1801(h) (2012).

Id. § 1805(a).

18 U.S.C. § 2703(a) (2012). An electronic communications system provides customers “the ability to send or receive wire or electronic communications.” Id. § 2510(15).

Id. § 2703(b). A remote computing service provides “computer storage or processing services by means of an electronic communications system.” Id. § 2711(2).
PREVENTING “GOING DARK”: A SOBER ANALYSIS AND...

158 Id. § 2703(d).


160 Id.


162 Frederic Lardinois, Gmail Now Has More Than 1B Monthly Active Users, TECHCRUNCH (Feb. 1, 2016), http://techcrunch.com/2016/02/01/gmail-now-has-more-than-1b-monthly-active-users/.

163 Id.

164 Frederic Lardinois, Gmail Now Has 900M Active Users, 75% On Mobile, TECHCRUNCH (May 28, 2015), https://techcrunch.com/2015/05/28/gmail-now-has-900m-active-users-75-on-mobile/.

165 ALEXANDER DE LUCA ET AL., EXPERT AND NON-EXPERT ATTITUDES TOWARDS (SECURE) INSTANT MESSAGINGG 147-51 (2016).

166 Id. at 149.

167 Id.

168 Id.

169 Id.

170 See Benjamin Wittes, Thoughts on Encryption and Going Dark, Part II: The Debate on the Merits, LAWFARE (July 12, 2015, 2:00 PM), https://www.lawfareblog.com/thoughtsonencryptionand-going-dark-part-ii-debate-merits (without surveillance capabilities, law enforcement would be largely unable to investigate illicit behavior in cyberspace, which would allow malicious actors to conduct activities unbridled).

171 CHERTOFF GROUP, supra note 1, at 5-6.


173 Id.

174 U.S. CONST. amend. V.


178 Id. at 1346-49. There is an exception to this rule when the government can demonstrate that the existence of the information sought is a “foregone conclusion.” In Fisher, the Supreme Court held that a testimony was a “foregone conclusion” and therefore did not implicate the Fifth Amendment’s protection against self-incrimination because the government could show
with “reasonable particularity” that it already knew the location, existence, and authenticity of the materials that it sought to compel. *Fisher*, 425 U.S. at 410-11.

179  *Schmerber v. California*, 384 U.S. 757, 764 (1966) (acknowledging that compulsion of physical evidence does not violate the Fifth Amendment).


181  *Id.* at 1447-48.

182  *Id.* at 1448.


184  *Id.* A legal order to facilitate surveillance issued in one country may not be satisfactory to an entity in another country. This would raise questions about which country's laws should take precedence, which country's court system would hear the dispute, and whether international agreements would be necessary to facilitate such cross-border orders and disputes.


186  *Id.*

187  *Id.*

188  *Id.*

189  *Id.*

190  *Id.*

191  *Id.*

192  *Id.*

193  *Id.*


196  *Id.*

197  *Id.*

198  *Id.*

199  *Id.*

200  *Id.*

201  *Id.*

Section 230(c)(1) states “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” *Id.*


*Id.*

28 UFLJLPP 65

REGULATING DOMESTIC INTELLIGENCE COLLECTION

Washington and Lee Law Review
Winter, 2014

Article
Emily Berman

Copyright (c) 2014 Washington and Lee University School of Law; Emily Berman

REGULATING DOMESTIC INTELLIGENCE COLLECTION

Abstract

Scholars have long recognized that a Federal Bureau of Investigation wielding robust domestic intelligence-collection powers poses a threat to civil liberties. Yet the FBI's post-9/11 mandate to prevent terrorist attacks (not merely investigate completed attacks) demands that the agency engage in broad intelligence-collection activities within the United States-activities that can threaten fundamental freedoms. This Article argues that strategies derived from administrative law principles can help alleviate the tendency of threat-prevention efforts to erode civil liberties.

The fundamental problem this Article tackles is that the traditional governance mechanisms we rely upon to protect individual rights are ineffective in the domestic intelligence-collection realm. This failure of traditional checks stems from, first, the absence of practical constraints to channel the enormous discretion that the Justice Department and the FBI enjoy in determining the scope and nature of the FBI's domestic intelligence-collection activities; second, the lack of judicial or political checks on these activities, resulting in a deficit of democratic legitimacy and accountability; and third, the risk that the FBI's singular focus on terrorism prevention will overwhelm rights-protection concerns.

Drawing on principles of administrative law, this Article explains how regulatory strategies can be employed to improve governance of domestic intelligence gathering. It recommends imposing procedural requirements on the exercise of discretion, facilitating meaningful pluralist input into relevant decision-making processes, and augmenting the attention given to civil liberties concerns by requiring the Justice Department to prepare Civil Liberties Impact Statements and by including in the process an entity whose primary goal is the protection of civil liberties. These governance reforms will prompt domestic intelligence regulation to take account of civil liberties while preserving the ability of law enforcement to pursue security.

Table of Contents

I. Introduction 5
II. The FBI's Domestic Intelligence-Collection Powers 12
   A. The Scope of the FBI's Intelligence-Collection Powers 16
   B. Existing Failures of Intelligence-Collection Governance 28
      1. Intelligence Collection and Practical Constraints 30
      2. Intelligence Collection and Judicial or Political Constraints 32
      3. Intelligence Collection and Structural Constraints 38
   III. Administrative Strategies as Governance 42
        A. Preliminary Questions 43
        B. Reason-Giving as Constraint 49
In the summer of 2013, former National Security Agency contractor and Central Intelligence Agency employee Edward Snowden shook the world with revelations of extensive United States government surveillance activities—including surveillance of American citizens. The revelations sparked a renewed debate about the proper scope of intelligence collection in a democracy. What has gone unaddressed in this debate, however, is the vast investigative powers conferred on America's domestic surveillance agency—the Federal Bureau of Investigation (FBI). And unlike the foreign-intelligence programs disclosed by Snowden's leaks, some of the FBI's most powerful domestic intelligence-collection authorities include neither statutory limits nor judicial oversight.

If the FBI determines that an individual's daily life is relevant to a terrorism investigation, it can easily draw a detailed picture of that life. With no reason at all for suspicion and no judicial approval, agents can follow the individual around the clock to ascertain where he goes. They can ask his neighbors about their conversations with him, or dispatch an informant to his house of worship to report on the individual's religious observance. With the minimal process associated with issuing an administrative subpoena, the government can establish a record of the individual's movements and social dealings by acquiring financial and employment records, a list of email addresses with which he has corresponded, and a list of phone numbers he has dialed.

These broad investigative powers operate in tension with fundamental rights. Collecting extensive personal information about innocent Americans raises concerns about privacy; about impact on freedoms of expression, association, and religious practice; and to the extent that such activity is disproportionately focused on particular communities, it can raise equal protection concerns as well. The challenge, then, is how to mitigate these civil liberties threats without unduly interfering with the FBI's ability to prevent terrorist attacks.

This Article argues that administrative law strategies suggest several measures that, taken together, would represent a domestic intelligence governance regime better equipped to safeguard civil liberties. This argument bridges a gap between two growing areas of literature. The first, which I label “risk-management literature,” advocates taking a regulatory approach to the threat of terrorism—treating it not as an enemy to defeat, but, like environmental or health and safety risks, as a chronic problem to be assessed and managed. The other, the “rights-protection literature,” argues that the current domestic intelligence-collection governance regime fails to address effectively the tension with civil liberties created by the FBI's contemporary counterterrorism efforts. I contend that the risk-management literature's valuable insight that administrative law can usefully be applied to improve governance in the security context suggests a means of mitigating the rights-protection scholars' concerns.
These rights-protection concerns arise in large part because traditional means of regulating executive power cannot effectively protect civil liberties in this area. This failure of traditional checks results from three characteristics of the regime regulating domestic intelligence collection. First, it lacks both doctrinal and practical constraints on the FBI and Justice Department's enormous discretion in drafting and implementing the applicable rules. 

Second, the checks that normally ensure the accountability and democratic legitimacy of government actions-judicial review, congressional oversight, and public scrutiny-simply do not operate effectively in the secretive world of intelligence collection. And third, in its vigorous pursuit of terrorism prevention, the FBI is not subject to any structural checks to prevent it from undervaluing rights protection.

This broad discretion, democracy deficit, and absence of counterweight to the FBI's prevention goal means that-despite the privacy and liberty implications of the FBI's activities-the responsibility of striking a balance between security needs and other important interests is left almost entirely to the Attorney General and the FBI itself. Restraints on the FBI's domestic intelligence-collection activities come from the Attorney General's Guidelines for Domestic FBI Operations (Guidelines) and the Domestic Investigations and Operations Guide (DIOG). The Guidelines are developed by the Attorney General and set out a basic framework for the FBI's operations; the several-hundred-page DIOG, promulgated by the FBI, specifies more detailed rules for the Guidelines' implementation. The Guidelines were originally created to forestall impending legislative efforts to cabin the Bureau's intelligence-collection powers but due to subsequent changes-particularly post-9/11 changes aimed at promoting terrorism-prevention efforts-the Guidelines and DIOG now serve to facilitate rather than limit the Bureau's intelligence-collection role.

Conceptualizing the threat of terrorism as a regulatory challenge represents a significant first step toward devising an improved governance regime for the Guidelines and the DIOG. But there are additional steps that the existing risk-management literature has not yet taken. The idea of how administrative law would apply to intelligence collection remains undertheorized-an unsurprising state of affairs given the recent emergence of this line of inquiry. Scholars have not yet tackled the thorny question of how to truly reap the benefits that the regulatory approach offers, while simultaneously taking into account the unique governance challenges presented by the domestic-intelligence context. Nor has the literature to date sufficiently grappled with civil liberties concerns.

This Article seeks to fill these gaps, drawing on administrative law principles to suggest novel governance designs custom tailored to address the civil liberties concerns inherent in domestic intelligence collection. To do so it looks to areas of administrative law that present similar governance challenges, identifies how those challenges have been addressed in the administrative state, and suggests how to adapt those strategies to function in the intelligence-collection context.

Each of the governance challenges this Article identifies has an analog in the administrative state. Take first the scope of discretion conferred on federal officials. This delegation of discretionary authority to the Attorney General and the FBI resembles the broad delegations in statutes establishing administrative agencies' powers and responsibilities. Concerns arising from the scope of these delegations are addressed through the Administrative Procedure Act's procedural requirements. Second, the absence of judicial, legislative, or public involvement in the design and implementation of the Guidelines is reminiscent of the democracy deficit inherent in the promulgation of regulations by technical experts so often at the heart of debates over the legitimacy of the administrative state. The administrative state has responded to this deficit by developing mechanisms to increase the participatory nature of administrative activity. And third, the risk that the Guidelines privilege the FBI's primary mission-the prevention of terrorism and protection of national security-over concerns for fundamental rights mirrors the many circumstances where agencies are charged simultaneously with multiple, competing goals. A menu of regulatory tools has been developed to reconcile competing
agency missions. Looking to the lessons that can be gleaned from these administrative governance strategies provides a useful roadmap for filling the intelligence-collection governance gap in a way that also protects civil liberties.

The Article proceeds in three parts. Part II first argues that the expansive scope of investigative authority that the current Attorney General's Guidelines and the FBI's implementing procedures confer on the Bureau exist in tension with the protection of civil liberties. It then contends that the lack of practical, traditional judicial or political, and structural checks to impose effective oversight on domestic intelligence gathering necessitates the implementation of alternative governance mechanisms.

In Part III, the Article harnesses administrative law strategies to suggest regulatory tools custom tailored to yield regulatory benefits that protect fundamental rights in the context of the FBI's Guidelines regime. To do so, it first explores the underlying purposes and justifications of traditional administrative law tools that are used to address governance challenges conceptually similar to those that the Guidelines regime presents. It then suggests specific reforms designed to achieve the same ends as those traditional tools in the intelligence-collection context. First, the Attorney General's discretion in the development and implementation of the Guidelines should be subject to a reason-giving requirement. Specific procedural limitations should require that the Attorney General or FBI Director provide both notice of a decision to amend the Guidelines or the DIOG and written justifications for their ultimate decisions. Second, to enable meaningful pluralist input into the process, the FBI must go beyond the cursory meetings with interested stakeholders that it has relied upon to date. Instead, a variety of entities inside the government should be empowered to participate in the amendment process. Third, to balance the government's interest in security with privacy and liberty concerns, (1) any changes to the Guidelines regime should be accompanied by a “Civil Liberties Impact Statement” prepared by the Justice Department, and (2) stakeholders whose primary goal is the protection of liberties- rather than the pursuit of security-should be involved in the process of drafting or modifying the Guidelines and their implementing regulations. This approach will impose a meaningful governance regime crafted for domestic intelligence gathering that mitigates concerns about the impact on civil liberties without sacrificing security.

Part IV addresses some possible sources of skepticism for this proposal. It first explains why reliance on existing administrative law tools-even if those tools are partially modified to operate in the intelligence-collection realm-will be insufficient to protect civil liberties. It then responds to questions about how to enforce these reforms in the absence of judicial review, suggesting alternative compliance mechanisms.

II. The FBI's Domestic Intelligence-Collection Powers

Since the FBI's inception, there has been tension embedded in its mission. It is charged not only with solving crimes but also with preventing them. While the two goals often complement one another, they call for very different types of investigative activities. Focus on crime solving argues for a set of investigative powers enabling inquiries into specific acts, with an eye toward successful prosecution of the perpetrators. Preventive work, by contrast, requires the collection of much broader swaths of information-information about illicit organizations, their members, their goals, their capacities, and their sources of funding as well as information about possible targets.

Over time, both the Bureau's focus and the rules governing its activities have swung back and forth along the spectrum between the targeted investigations of crime solving and the broader intelligence gathering associated with prevention. The Guidelines themselves are the product of the FBI's early-1970s move away from intelligence collection. After the United States Senate Select Committee to Study Governmental Operations with Respect to Intelligence Activities,
commonly known as the Church Committee for its chair Senator Frank Church (D-ID), *13 revealed that decades of unregulated intelligence collection by the FBI had resulted in widespread abuses of the government's investigative powers, 29 Congress determined that the FBI should be subject to a legislative charter setting out strict limits on its intelligence-collection authority. 30 In an effort to stave off potentially more restrictive legislative action, President Gerald Ford's Attorney General, Edward Levi, issued in 1976 the first set of Attorney General's Guidelines-known as the Levi Guidelines. 31

The Levi Guidelines strictly curtailed domestic intelligence investigations through a basic regulatory structure that subsequent versions of the Guidelines have largely retained. 32 *14 This structure consists of multiple investigative levels. For each successive level, a higher threshold of suspicion is necessary to proceed; the investigative tools agents may use are more intrusive; and procedural safeguards, such as the need for supervisory approval and limits on the temporal length of investigations, are more robust. 33 The Guidelines continue to function as the primary constraint on the FBI's operations and remain a justification for the lack of a statutory charter governing the FBI's activities, but they have not remained static. 34 Multiple modifications made in the years between 1976 and 2001 eased, though ultimately retained, restrictions on intelligence collection. 35

With 9/11, however, came a wholesale rejection of the anti-intelligence-collection mindset of the Levi era, resulting in a dramatic shift in favor of an aggressive prevention paradigm. The FBI's prioritization of preventing terrorism was reflected not only in the allocation of its resources, its focus, and its conception of its core mission but also in some dramatic modifications to the Guidelines themselves. 36 The eventual result was a set of *15 Guidelines-which were implemented by Attorney General Michael Mukasey in 2008 and remain in effect today-embODYING an unprecedented license for domestic intelligence collection and delegating to the FBI responsibility for imposing limits on that power. 37

In this Part, Section A demonstrates the scope of some of the FBI's contemporary intelligence-collection powers and their potential to create tension with privacy; rights of association, expression, and religious exercise; and equal protection principles. 38 Section B then argues that the absence of nondoctrinal checks-whether practical, judicial, political, or *16 structural-on domestic intelligence collection means that the only constraints on the FBI's intelligence-collection powers are the internal rules that the Justice Department and the FBI have imposed on themselves.

A. The Scope of the FBI's Intelligence-Collection Powers

Statutory and constitutional doctrine provides very few limits on government access to a vast amount of information about innocent Americans. 39 Any information that we have disclosed to a third-party individual or business entity, for example, lacks Fourth Amendment protection against unreasonable searches and seizures. 40 Thus the Constitution places no limits on the collection of information contained in credit card transactions, bank records, Internet service provider (ISP) records, Amazon.com transaction histories, Facebook activities, electronic toll records, cell-tower location data (in some jurisdictions 41 ), and even statements made to undercover agents or government informants-regardless of whether the agent or informant discloses his intention to share the contents of the conversation. 42 The First Amendment similarly lacks purchase *17 here. Intelligence-collection powers that impact religious practice would likely run afoul of the Free Exercise Clause as a facial matter only if they were being implemented with the purpose of suppressing religious exercise. 43 And if the Guidelines chill expression or curtail association, the activities they permit would be immune to facial constitutional challenge so long as they were narrowly tailored to further the compelling
interesting in preventing terrorism, an interest that has been afforded great weight by the courts. Statutes provide only slightly more protection.

The permissiveness of the doctrine means that the FBI's intelligence-collection powers face very few external legal constraints. In this doctrinal vacuum, the Attorney General was able to make several post-9/11 amendments to the Guidelines that facilitate an aggressive intelligence-collection role for the FBI. The first relevant amendment is the Guidelines' expression of the FBI's newly adopted preventive mission. Specific language explicitly affirms the FBI's role in the intelligence community and specifies authority to collect, retain, and analyze information for intelligence purposes. The Guidelines declare that “[t]he FBI is an intelligence agency as well as a law enforcement agency . . . [whose] functions accordingly extend beyond limited investigations of discrete matters” and urge the Bureau to use its analytic authority to “identify and understand trends, causes, and potential indicia of criminal activity and other threats to the United States that would not be apparent from the investigation of discrete matters alone.” To facilitate this analytical project, the Guidelines provide that all information collected “at all stages of investigative activity is . . . to be retained and disseminated for [intelligence purposes to facilitate the solution and prevention of crime, protect the national security, and further foreign intelligence objectives] regardless of whether it furthers investigative objectives in a narrower or more immediate sense.” Even information that wholly exonerates a group or individual from suspicion remains in government databases for storage, analysis (sometimes by algorithmic data-mining), and dissemination for inclusion in other government agencies' databases.

The 2008 Mukasey Guidelines also expanded the Bureau's collection powers to further its preventive mission—both with respect to what information it is permitted to collect and what tactics it may employ in that collection. The most significant of these expansions is the authorization of “assessments”—a new investigative stage. Assessments, which are inquiries designed to determine whether further investigation is warranted, require only an “authorized purpose,” meaning that the FBI must merely determine that it is acting to protect against criminal or national-security threats, or to collect foreign intelligence. There is no need for any concrete facts, evidence, or reason to believe that the subject of an assessment is involved in criminal or threatening activity. In other words, assessments may be undertaken in the absence of any “factual predicate.” Until 2008, some form of factual predication was required to initiate any level of investigation. The introduction of nonpredicated investigations is thus a significant expansion of the FBI's power.

Despite the absence of the need to establish a factual basis for an assessment, agents conducting assessments under the new Guidelines have available to them a wide array of highly intrusive investigative tools. During an assessment, the FBI may: (1) recruit and task informants to attend surreptitiously First Amendment-protected gatherings—such as religious services or political demonstrations—to collect information about what takes place there and who attends; (2) engage people in conversation while misrepresenting the agent's status as a federal official (so-called “pretext interviews”)—such as posing as an investigative target's new neighbor or business associate in order to gather information about her from friends, neighbors, and colleagues; (3) station agents outside a target's home or office—or even have them followed—so that their movements are tracked day and night; and (4) search commercial online services and resources, records maintained by “other federal, state, local, or tribal, or foreign governmental entities or agencies,” and all FBI and Justice Department records. This includes the FBI's National Security Branch's data set “[c]omposed of government information, commercial databases and records acquired in criminal and terrorism probes” that includes international travel records of citizens and aliens; financial forms; hotel and rental car records; and credit card transaction records. The government can not only search these databases for particular information, but also use them to perform
analysis based on a “pattern of behavior and search for that pattern in data sets.” Ever-more sophisticated data-mining tools render a detailed dossier on any American—even one entirely above suspicion—just one mouse click away. Thus, the Guidelines now permit, with no factual predicate, tactics that before 9/11 had been reserved for investigative stages whose initiation required at least some relevant evidence.

*21 Even investigative tactics statutorily limited to instances in which there is a factual basis for suspicion may be permitted based on very tenuous links to suspected wrongdoing. Agents can issue National Security Letters (NSLs)—a form of administrative subpoena—to access without a court order certain information about individuals that is “relevant” to a terrorism investigation. The relevance standard means that the individual about whom the FBI seeks information—internet or telephone subscriber information (possibly including cell-phone-generated location data providing a minute-by-minute account of an individual’s movements), internet search records, financial records—need not herself be a target of an investigation. And after acquiring a court order under §215 of the PATRIOT Act, the FBI may demand an even broader swath of information—“any tangible thing” that is “relevant to” “a full investigation,” a definition the government has interpreted to include the authority to collect all noncontent data regarding phone calls into, out of, and within the United States.

*22 In addition to the obvious privacy concerns raised by these rules, Professor Daniel Solove has also pointed out the potential for chill to First Amendment-protected activity when the FBI is engaged in widespread intelligence-gathering activities such as acquisition of internet search history; book purchases; phone call or email records; banking records; questioning an individual’s friends, neighbors, or colleagues; infiltrating religious or political gatherings; or using ISPs to identify individuals writing anonymous political blogs. Indeed, there is evidence that the FBI’s tactics have had this very effect in Muslim communities, where there is evidence of chill on attendance at political demonstrations, donations to political causes, speaking out against U.S. foreign policy, participating in community organizations, internet use, and book purchase habits.

Law enforcement surveillance of antiwar protesters and other political dissenters raises similar concerns about expressive and associational activities. A 2010 Justice Department review found that investigations of Greenpeace, People for the Ethical Treatment of Animals, and antiwar groups used “troubling” tactics. In the course of these investigations, the FBI classified nonviolent civil disobedience as “acts of terrorism,” extended investigations “without adequate basis,” and unnecessarily placed several Greenpeace members on federal watch lists. Members of similarly situated groups may harbor concerns that their political activities will attract government attention, the result of which could be inclusion on a watch-list or even becoming the target of an investigation.

The FBI’s aggressive investigation and infiltration of houses of worship can also chill the exercise of religious freedoms. As a result of the FBI’s use of undercover agents or informants in mosques, Muslim leaders across the country have reported “a reduction in attendance at mosques, a change in the language used at worship services, a decrease in contributions to Muslim charities, and an erosion of the trust and good will that are essential to the vitality of a religious community.” Studies have also shown that these tactics have also deterred members of the Muslim community from wearing clothing that expresses religious or cultural identities.

*24 Any of these intelligence-collection tactics may also be discriminatorily implemented, disproportionately burdening particular minorities. The use of unpredicated investigations in particular opens the door to investigative decisions based on the use of race, ethnicity, national origin, or religion. Historically, when law enforcement officials have been permitted
to collect intelligence on groups and individuals suspected without any objective basis of harboring ill will toward the United States, the burden of that investigative activity has fallen on groups that espouse disfavored ideologies, minorities, or others who are perceived as threatening. 75

The more specific DIOG rules magnify concerns about discriminatory implementation. Under the DIOG, the FBI may collect information regarding ethnic and racial behaviors “reasonably believed to be associated with a particular criminal or terrorist element of an ethnic community.” 76 Official expressions regarding how individuals become “radicalized” actually identify certain religious activities—such as Muslim men growing beards—as potential precursors to violent extremism, thus providing theoretical grounds for investigative decisions based at least in part on exactly these types of activities. 77

Investigative activity based in part on racial or ethnic characteristics is not limited to scrutiny of individuals. FBI policy also embraces “community mapping”—the practice of collecting and storing information about particular ethnic communities. 78 *25 The rules “permit the FBI to identify locations of concentrated ethnic communities” as well as to collect “the locations of ethnic-oriented businesses and other facilities” (including religious facilities such as mosques) because “members of certain terrorist organizations live and operate primarily within a certain concentrated community of the same ethnicity.” 79 The DIOG justifies such activity by asserting that concentrations of certain ethnic communities provide an opportunity for “identified terrorist subjects from certain countries [to] relocate to blend in and avoid detection.” 80 Under these authorities, the FBI has collected information about religious, ethnic, and national-origin characteristics of American communities, identifying “Arab-American and Muslim communities in Michigan as a potential terrorist recruitment ground” as well as noting “an increase in the African-American population of Georgia when analyzing *26 ‘Black Separatist’ groups,” pointing to Chinese and Russian communities in San Francisco “as a place to look for organized crime syndicates,” and “highlighted Latino communities as potentially harboring the Central American gang MS-13.” 81

The Guidelines are not blind to the concerns these tactics raise. In fact, they bar the FBI from initiating investigations into U.S. persons “solely for the purpose of monitoring activities protected by the First Amendment.” 82 Nor may the Bureau “predicate an investigation simply based on somebody's race.” 83 But the Guidelines as well as the DIOG prevent such activity only when it is motivated solely by the desire to monitor First Amendment-protected activities, or by race, religion, or national origin. 84 Investigative activity prompted in part by these factors is not barred. 85 Individuals thus can be singled out for scrutiny due, at least in part, to their political or religious expressions, *27 activities, or associations. This profiling on the basis of “national origin plus” could expose a large population of innocent persons to FBI scrutiny. Indeed, as law enforcement officials told the Associated Press, “[a]mong the factors that could make someone the subject of an investigation is travel to regions of the world known for terrorist activity . . . along with the person's race or ethnicity.” 86 Thus, every individual of Pakistani origin who travels to Pakistan to visit family is conceivably at risk of being subjected to FBI investigation merely on that basis.

A final notable modification to the Guidelines implemented in 2008 was the elimination of the vast majority of oversight provisions contained in prior iterations of the Guidelines—time limits on investigations, the need to obtain supervisory approval, requirements to report regularly to FBI Headquarters or the Justice Department. 87 Instead, these restrictions have been relegated to the DIOG, thereby empowering the FBI to determine the scope of its own power in this regard. 88 Indeed, the DIOG may be changed whenever the FBI—not the Attorney General—determines that it should be (as it was in 2011 89), and FBI *28 leadership can authorize departures from the DIOG’s requirements. 90
There may have been good reason for some of these changes. Indeed, some of them were the result of recommendations made by the 9/11 Commission to improve America's counterterrorism capacity. There have, after all, been significant changes since 1976—in the threats that we face, in the need for intelligence gathering, in Americans' expectations of what their government should do to protect them and their interests. The limits on FBI intelligence activities imposed in the 1970s, however, reflected concerns expressed by the legislature and the public. The current Guidelines, by contrast, have fundamentally transformed the role of America's primary domestic federal law enforcement agency with almost no public debate and with no legislative action. Thus, regardless of what one thinks about the propriety of the changes themselves, we might question the adequacy of the process leading to such a transformation.

B. Existing Failures of Intelligence-Collection Governance

If the FBI's intelligence-collection authorities do not run afoul of existing legal limits, why is the way in which they are governed a cause for concern? Because despite these authorities' undisputed implications for civil liberties—indeed, their tendency to result in civil liberties infringements is what inspired the implementation of the Guidelines in the first instance—they are untouched by the nondoctrinal constraints that usually accompany law enforcement activities. This means that the only constraints on the FBI's intelligence-collection powers are the internal rules the Justice Department and the FBI have imposed on themselves.

Several inherent differences between intelligence collection and crime-solving investigations account for the inapplicability of constraints that usually limit government action. First, the very nature of the intelligence-collection enterprise is inherently more expansive in scope—proactive rather than reactive and less narrowly targeted. Consequently, practical constraints that usually serve to limit law enforcement agencies' activities-resource limitations and a focus on solving individual crimes—are inapplicable. Second, the secretive nature of intelligence-collection activities renders them effectively immune to judicial review as well as scrutiny from the legislature and the public. And third, the Justice Department and the FBI generate the Guidelines and the DIOG in the context of the FBI's post-9/11 focus on terrorism prevention. This means that the rules are crafted by government officials with security and intelligence-collection expertise. There are, therefore, no structural checks to remove from the hands of security technocrats the normative judgments that must be made about the relative importance of aggressive intelligence collection and rights protections. The result is that, despite the fact that intelligence collection implicates important values, the only constraints on that collection are effectively self-imposed.

1. Intelligence Collection and Practical Constraints

Intelligence collection compliments and overlaps with criminal investigation, but it is a distinct endeavor. Crime-solving efforts are tied to individual cases. They focus on the investigations of specific acts in an effort to collect evidence related to each element of a completed or impending crime, and tend to end with a decision to prosecute or not to prosecute. As previous versions of the Guidelines recognize, criminal investigations are more circumscribed in scope and tend to be shorter in duration than intelligence investigations. Intelligence investigations, by contrast, call for much different—and much broader-investigative activities, which “may continue for several years.” Furthermore, the focus of such investigations “may be less precise than that directed against more conventional types of crime. . . . For this reason the investigation is broader and less discriminate than usual,” seeking information about potential targets as well as criminal or terrorist organizations, their members, their goals, and their sources of funding. Thus, the
goal of intelligence collection is to gather as much information as possible for future analysis, rather than seeking only information connected to a discrete incident.

*31 A law enforcement agency focused on crime solving is less likely to allocate time and resources to the types of activities currently taking place under the Guidelines. While that agency may be permitted, for example, to place an individual under twenty-four-hour surveillance or attend political rallies, it is unlikely to devote scarce manpower to such activities if there is no factual basis for doing so. If an agency's primary goal, however, is to collect as much information as possible to include in a database because it is impossible to know what information might eventually lead to the prevention of a threat or crime, then its ability to engage in that collection will be supported by sufficient funds and manpower. Recent technological advances exacerbate this phenomenon because they have made the collection and storage of information infinitely cheaper and easier than ever before.

In order to ensure that the FBI engages in the broad investigative activities associated with intelligence collection, the Guidelines—initially imposed to restrict domestic intelligence collection—have been transformed into Guidelines expressly facilitating and encouraging such activity. They remind agents that they “cannot be content to wait for leads to come through the action of others” and thus “must proactively draw on available sources of information.” In so doing, they both expand the FBI’s authorized powers and relax the limits on how those powers may be used. Because they will be utilized far more expansively than their crime-solving counterparts, which existing doctrine evolved to regulate, the FBI’s powers present a greater threat to fundamental values than existing doctrine may indicate.

2. Intelligence Collection and Judicial or Political Constraints

In engaging in the broad intelligence-collection activities envisioned by the Guidelines, the FBI will elude traditional checks on power, such as judicial review and congressional or public oversight. The result is that the Guidelines and their implementation lack both the accountability and the democratic legitimacy that usually accompanies government policy.

There are several obstacles to judicial review of the Guidelines and the activities undertaken pursuant to them. As an initial matter, the Guidelines themselves disclaim any intention to create enforceable rights, so any action taken pursuant to them can be challenged only if it is otherwise unlawful. In addition, the secrecy of these activities ensures that individuals who seek to challenge intelligence-collection regimes will struggle to demonstrate a sufficiently concrete injury to establish standing to sue. Surveillance tactics are designed to prevent targets from being alerted to the fact that law enforcement is gathering information about them, so it is difficult to point to specific government action causing harm. Moreover, courts have held that neither allegations of general chill nor an “objectively reasonable likelihood” that a plaintiffs' communications will be subject to surveillance are sufficient. Thus, standing remains a bar to the courthouse door.

Another barrier that has proved fatal to judicial review of intelligence collection is the state secrets privilege, which allows the government to withhold evidence whose disclosure might endanger national security. At times the privilege results in a case being dismissed outright. In other instances, a suit may proceed with evidence that the government is willing to share only ex parte, undermining the proceedings' adversarial nature.
The government's investigative actions are most frequently scrutinized through motions to suppress evidence collected in violation of the Fourth Amendment. But this process generally eludes the targets of surveillance. When the government gathers information for the purposes of criminal prosecution and seeks to introduce it as evidence at trial, only the individual about whom the information was gathered—the criminal defendant—will have the opportunity to challenge the government's actions through a suppression motion. This means that these practices will face challenges in those circumstances where the government's case is most compelling—when a guilty person seeks to exclude probative inculpatory evidence. Moreover, much of the government's intelligence-collection activity never leads to prosecution. As a result, innocent targets of surveillance—those whose information is collected because it is deemed “relevant” to an investigation, or members of a house of worship who change their religious practices due to fear of surveillance—will be unable to invoke judicial protection.

The accountability gap left by the absence of judicial review will not be filled by legislative or public scrutiny. The origin story of the original Attorney General Guidelines included a significant role for Congress. Having been prompted by the Church Committee's findings of misconduct and the resulting public outrage, the legislature was intimately involved in developing the contents of the Guidelines. Congress held a series of hearings on the issue over the course of several years, and saw its suggestions ultimately reflected in the Levi Guidelines. Given that the Guidelines were implemented, at least in part, to avoid more stringent legislative action, this is not a surprise. Surely Attorney General Levi knew that if the rules he instituted did not appear to address Congress's concerns, they would fail to sap the momentum for enacting a statutory charter for the FBI.

The contemporary political economy of congressional oversight in this area means that legislative oversight will not provide any more effective a check than judicial action. Legislators' incentives weigh against aggressive involvement. The downside risks of unsuccessful counterterrorism policies (additional attacks) are high. If those policies are developed outside of the legislative process, Congress can share (if not entirely evade) blame. Moreover, counterterrorism policy “is a subject matter that is especially prone to legislative delegation because it often entails hard trade-offs,” which are the types of questions Congress is least likely to address. In addition to undermining legislative involvement in counterterrorism policy formulation, existing institutional features also render congressional oversight of domestic intelligence-collection policy ineffectual. Congress, of course, retains oversight authority over the FBI. If it wants to play a more active role in overseeing the Guidelines, it has the tools to do so. After all, Congress determines whether and to what degree the FBI's intelligence-collection activities are funded. Moreover, the relevant committees of jurisdiction conduct regular oversight hearings at which the Attorney General and FBI Director appear. Legislators can ask Justice Department and FBI officials for information about the Guidelines or the FBI's activities at any time.

Perhaps as a result of the existing incentive structure, however, Congress has shown little appetite to pursue Guidelines-related issues of late. The most recent modification to the Guidelines, for example, failed to reflect congressional input. The Justice Department provided the Senate Judiciary Committee a completed draft of the Mukasey Guidelines a few months before they were implemented. A handful of senators requested that Attorney General Mukasey delay their implementation until Congress had the opportunity to develop suggestions regarding ways to minimize civil liberties concerns. Their request went unanswered. And even when FBI Director Robert Mueller III inaccurately testified in 2010 before Congress that the FBI did not have the authority to conduct unpredicated investigations, legislators took no follow-up action.
While Congress has shown little interest in scrutinizing the Guidelines, the public is not given a choice in the matter. Activity undertaken pursuant to the Guidelines is secret and therefore rarely apparent on its own or reported in sufficient detail in the news media. Moreover, information about how the Guidelines are used is exempt from disclosure under the Freedom of Information Act on the basis of either the law enforcement or the classified-information exemption contained in that statute. The public also lacks means to scrutinize how the rules are implemented. In some cases even the rules themselves are secret. The publicly available version of the DIOG, for example, entirely redacts the rules governing undisclosed participation in religious or political gatherings.

3. Intelligence Collection and Structural Constraints

The governance concerns posed by the FBI's intelligence-focused mission and the lack of traditional oversight are compounded by the FBI's tendency to emphasize that mission over other concerns. The Attorney General and the FBI are responsible for incorporating two sometimes conflicting responsibilities into intelligence-collection policies—terrorism prevention and civil liberties protection. Given the incentives to err on the side of security, the Guidelines risk short-changing civil liberties concerns when the two missions conflict. The FBI declares on its website that “as an intelligence-driven and a threat-focused national security organization with both intelligence and law enforcement responsibilities,” its mission is to “to protect and defend the United States against terrorist and foreign intelligence threats, to uphold and enforce the criminal laws of the United States, and to provide leadership and criminal justice services to federal, state, municipal, and international agencies and partners.” Similarly, the Bureau identifies as its top priority protecting the United States from terrorist attack, followed by combatting foreign intelligence operations, cyberattacks, high-technology crimes, and public corruption.

In other words, despite its location in the Department of Justice and its law-enforcement responsibilities, the FBI is now primarily a national security and intelligence-focused agency. In all of the FBI's statements listing its goals and priorities, protecting civil liberties falls far below protecting against terrorist attacks and other security threats. Like the FBI's other statements of its priorities, the post-9/11 Guidelines themselves present “[p]rotection of the United States and Its people” as the goal that the FBI's investigative authorities are designed to further. And while the Guidelines have always explicitly required investigators to use the least intrusive method possible to achieve their investigative goals, as of 2002 the Guidelines include a caveat: agents should not “hesitate to use any lawful techniques consistent with these Guidelines, even if intrusive,” where the degree of intrusiveness is warranted in light of the seriousness of a threat, or in light of the importance of foreign intelligence sought in the United States' interests. This point is to be observed in particular, agents are instructed, in investigations relating to terrorism.

The DIOG presents even larger concerns on this front. Those rules are devised and implemented within the FBI itself. Thus, any nonsecurity perspective that Justice Department officials may bring is absent from the DIOG. And because the terms of the Guidelines are relatively skeletal, the FBI is left to fill in most of the details itself. The DIOG does include hortatory language regarding the need to be solicitous of civil liberties; to refrain from infringing on First Amendment rights and from profiling on the basis of race, religion, ethnicity, or national origin; and to limit appropriately the scope and intrusiveness of FBI activity. It also instructs that “when First Amendment rights are at stake, the choice and use of investigative methods should be focused in a manner that minimizes potential infringement of those rights.” But after warning agents to take privacy, equal protection, and First Amendment rights into account, the DIOG concludes
by reiterating that “FBI employees may use any lawful method allowed, even if intrusive, where the intrusiveness is warranted by the threat to the national security or to potential victims of crime.”

A recent quote from former Director of National Intelligence Mike McConnell captures the sentiment of many security policymakers. In his view, the American people have “very little appreciation for the threat,” and “special interests, particularly civil liberty groups with privacy concerns,” prevent the intelligence community from doing its job as well as it otherwise could. This view of the need to consider privacy concerns as a hurdle to effective policy, rather than as an integral part of policymaking, illustrates the tension that sometimes arises between the FBI’s primary mission and its responsibility to protect civil liberties. As others have pointed out, the FBI will value success in carrying out its primary mission and will favor terrorism prevention “over competing concerns such as the protection of civil liberties.”

Pointing out the elevation of the FBI’s anti-terrorism mission over other considerations is not meant to be an indictment. That mission is a vital one that should be pursued vigorously. And with over a century of experience conducting criminal and security investigations, the FBI is the agency in the best position to determine the most effective means of pursuing that mission. This includes decisions regarding which investigative methods will be most successful in countering threats to the country. This expertise should not be undervalued.

At the same time, the decision about what level of intrusiveness society is prepared to accept in pursuit of security is not a matter of technical, investigative, or intelligence-collection expertise. Determining the intrusiveness of an investigation justified by any particular set of circumstances necessarily involves normative judgments implicating fundamental values. As should now be clear however, the only true constraints on the FBI’s intelligence-collection activities are the Guidelines and the DIOG. This leaves decisions regarding the appropriate balance between the FBI’s security mission and the interests on the other side of the scale in the hands of the Attorney General. He, in turn, has delegated many of those decisions to the FBI itself. Thus, concerns over security will have a prominent role in such decision making and other interests will be short-changed.

III. Administrative Strategies as Governance

If the usual tools of governance fail to apply effectively to the Guidelines and the DIOG, how do we devise mechanisms to fill this governance gap? This Part argues that we can borrow from the institutional design principles of the administrative state to address three specific challenges presented by the Guidelines regime. First, the absence of both doctrinal and practical limits on the FBI’s intelligence collection confers expansive discretion on the Attorney General and the FBI. Second, the lack of judicial, legislative, or public scrutiny of FBI policy results in a deficit of both accountability and democratic legitimacy. And third, the FBI’s focus on threat prevention creates a risk that its rulemaking decisions will give insufficient attention to liberty and privacy interests. With respect to each of these challenges, this Part identifies strategies the administrative state employs, and uses those strategies to develop concrete suggestions to improve intelligence-collection governance.

This examination of administrative law strategies-designed to channel discretion, increase accountability and legitimacy, and ensure that competing priorities are afforded sufficient attention-suggests the following concrete reform proposals. First, a reason-giving framework should be implemented that (1) requires the Attorney General to provide notice of his or her intention to modify the Guidelines; and (2) specifies that any modifications must be justified in writing. Second, to promote meaningful pluralist input, the Attorney General should be obligated to consider as part of the Guidelines-development process the views of stakeholders outside the intelligence community (though not necessarily outside the government). Third, in order to ensure liberty interests are not marginalized, (1) the Justice Department should...
be required to prepare a statement indicating the likely impact on civil liberties of any changes to the Guidelines, and
(2) the Privacy and Civil Liberties Oversight Board should be empowered to participate meaningfully in the Guidelines’
development. 157 Before beginning the analysis that suggests these particular reforms, however, a few preliminary points
are in order.

A. Preliminary Questions

This subpart will preemptively address several questions raised by the discussion that follows. First, it acknowledges
(and rejects) concerns, based in political realities, that these proposals can be no more than a thought exercise. Then it
clarifies the role that the Administrative Procedure Act plays in the argument, and finally it recognizes that the addition of
procedural requirements may impose costs as well as benefits. With these first-order questions addressed (if not entirely
resolved), the Article turns to the recommendations themselves.

First, some brief thoughts on political economy. This Article aims to propose some plausible reforms in an area where
what Professor Heather Gerken calls the “here to there” problem is a significant obstacle. 158 Perhaps even more than
in other policy areas, expectations that Congress will act to implement these recommendations-through legislation or
through other available levers of power—are likely to be disappointed. Indeed, congressional oversight of national security
policy has long been considered ineffective by government officials, outside task forces, and scholars. 159 The dearth of
public information about national *44 security policy, which makes oversight significantly more challenging, is partially
to blame. 160 But there are also perverse incentives at work: legislators have no incentive to engage in aggressive oversight
of intelligence-collection powers. 161 Legislators gain little by taking ownership over security policy. 162 Meanwhile, so
long as Congress can label such policies “executive,” it cannot be blamed for intelligence failures. 163 The result is that
all electoral incentives point toward congressional deference to executive policy preferences in this area. 164 This is *45
especially so for intelligence-collection policies imposing disproportionate impact on certain segments of society, such
as minorities or noncitizens, whose interests carry little electoral weight with legislators. 165 Expectations that Congress
will take action in this area are thus likely to be disappointed.

And if Congress is so impotent with respect to the oversight of the Guidelines, what would prompt either Congress or
the executive branch to impose the types of restrictions proposed here? While I do not want to minimize the challenges of
prompting government actors to impose restrictions on domestic security measures, opportunities for reform do arise.
My hope is to generate a menu of possible options, so that when such an opportunity presents itself, it may be exploited.

These opportunities will likely arise out of one of two possible exogenous events. First, there could be an event
that triggers widespread public concern about the FBI's activities and places sufficient pressure on Congress, or the
Attorney General, or the President that they have to take some action. After all, the revelations about the activities
of COINTELPRO are exactly what brought the Attorney General's Guidelines into existence. 166 And the recent
revelations regarding the NSA promise to prompt a series of policy changes. The second possible event is a judicial
decision invalidating a particular FBI policy. Such a decision is more likely to lead to modifications to particular
practices—such as we saw when a court rejected as unconstitutional the provision that barred National Security Letter
recipients from disclosing to anyone that they had received one 167—rather than to a broad procedural framework. But
like a scandal, judicial invalidation of certain FBI activities could spur a broader reform effort.

*46 Given the appropriate political environment, there are at least three reasons to think that the imposition of a
framework like the one suggested here is not entirely implausible. As an initial matter, there is the FBI’s concern over
legitimacy. The Bureau's ability to succeed in its mission requires constructive relationships with the communities in which it operates. Yet its aggressive intelligence-collection tactics—and their concentration in Muslim communities—has alienated many members of that community, raised suspicion and distrust of the Bureau in some quarters, and undermined cooperative relationships. Improved governance is thus not the only benefit that would flow from implementing APA-like procedures; institutionalizing rulemaking procedures would also yield improvements in community relations, public perceptions of legitimacy, and consequently, FBI effectiveness. In addition, government documents and scholarly commentary are replete with arguments about the value of process in legitimating government action. The FBI's practice of reaching out to nongovernmental organizations in anticipation of issuing new intelligence-collection rules indicates an awareness of the benefits of generating the support of outside stakeholders. Subjecting itself to a set of procedural rules would go far in this regard. And finally, none of the proposals here are substantive. They do not call upon the FBI to cede any particular powers, or to discontinue existing policy. Indeed, they acknowledge the Attorney General's and FBI's role in generating the rules by which the FBI operates, so long as they can show both that changes in the FBI's authority are needed and that the proposed changes are reasonable ones. And finally, when it comes to imposing limits on government actors, broad procedural frameworks often face less opposition than substantive policy changes.

A second preliminary note concerns the role of the APA. While several of the proposed reforms are inspired by provisions of the APA, this Article does not argue that the APA's procedural rules apply to the FBI as a matter of binding law. In fact, it does not take a position with respect to whether the Guidelines or the DIOG constitute legislative rules subject to APA requirements, or whether they represent informal guidance documents or “rules of agency organization, procedure, or practice,” which are explicitly exempt from many of the APA's constraints. Instead, the Article looks to the way the APA and other sources of administrative law address particular concerns and argues that intelligence-collection governance would benefit from implementing procedures inspired the animating principles behind these sources of administrative law.

The idea of imposing a governance framework on the development of rules in the absence of a statutory requirement to do so is not a novel one. The Office of Management and Budget (OMB)—part of the Executive Office of the President tasked with overseeing the regulatory decisions of administrative agencies—in its Final Bulletin for Agency Good Guidance Practices “establishes policies and procedures for the development, issuance, and use of significant guidance documents.” These policies are designed to increase the quality, transparency, consistency, fairness, and accountability of agency guidance practices. To that end, the Bulletin suggests that agencies engage in “procedures similar to APA notice-and-comment requirements” for some types of guidance documents in order to “increase the quality of the guidance and provide for greater public confidence in and acceptance of the ultimate agency judgments.” Similar language appears in the Department of Homeland Security's (DHS) official guidance on preparing Privacy Impact Assessments (PIAs), which all agencies—including the FBI—must generate for any substantially revised or new Information Technology System that collects, maintains, or disseminates personally identifiable information from or about members of the public. According to the DHS Guidance, requiring agencies to follow procedures designed to call attention to issues of legitimacy “demonstrates to the public and to Congress” that the new systems “have consciously incorporated privacy protections.” In other words, both OMB and DHS policy takes the position that procedural constraints result in both better substantive rules and an increase in the perceived legitimacy of those rules, even when those constraints are self-imposed rather than statutorily required.
A final preliminary note: implementing these reforms would not be costless. As an initial matter, any increase in the onerousness of modifying the Guidelines creates pressure to shift *49 policy-making decisions to a level where these rules do not apply, making accountability even more elusive. For this reason, any efforts at reform would have to apply to changes in the DIOG as well as the Guidelines, and consider ways to prevent further devolution of decision-making responsibility. These suggested changes would also, of course, consume time and personnel not currently devoted to the Guidelines. But these costs need not be prohibitive. As an initial matter, the costs themselves would impose a potentially valuable barrier to arbitrary or unnecessary changes. Only when changes are in fact necessary will the Attorney General or FBI Director undertake the amendment process. Moreover, the Guidelines and DIOG are modified so infrequently that the need to allocate additional resources to the project would be rare. If implemented effectively, these rare additional costs would be justified by their benefits.

B. Reason-Giving as Constraint

 Courts and commentators have raised a litany of reasons why extending broad discretion to administrative agencies can be problematic from a governance standpoint—reasons that apply with equal force to the FBI’s exercise of intelligence-collection powers. Agency strategy for channeling discretion, largely dominated by reason-giving requirements, is therefore an important source of ideas for addressing that concern in the context of the Guidelines regime.

1. The Downsides of Discretion

Consigning significant policy choices to administrative agencies operating with broad discretion undermines the constitutional mechanism of promoting both accountability and sound decision making. When it comes to legislation, the Constitution seeks to avoid these concerns by subjecting legislative decisions to the deliberation and contestation that serves as a bulwark against faction and tyranny. 182 Freed from *50 the requirements of Article 1, § 7, 183 however, agency decision makers might engage in a wise and thoughtful decision-making process; but they are equally capable of making poor choices, opting for policies that are uninformed, arbitrary, irrational, self-interested, or otherwise untethered to the public interest. 184 Absent some alternative check on the way in which discretion is exercised, there is therefore no reason to expect an agency’s decision-making process to result in the best outcome—however that is defined.

Similarly, because granting decision-making authority to bureaucrats not subject to electoral forces that constrain other policymakers removes those decisions from the field of political battle, Congress both eludes responsibility for making difficult policymaking decisions and insulates the policies themselves from electoral backlash. 185 Broad agency discretion thus undermines the very nature of participatory democracy and raises concerns about political accountability for critical decisions of national policy. 186

*51 And while the Supreme Court’s decisions limiting legislative delegations to agencies were confined to the New Deal-era, 187 so long as Congress sets down an “intelligible principle” for the agency to follow, 188 many of the procedural rules developed in the administrative state serve to cabin discretion. 189 Thus, while agency decision makers continue to enjoy significant leeway, the threat to democracy and accountability posed by agency discretion has not gone unaddressed.

2. Channeling Discretion into Reasoned Decision-Making

The administrative state has grappled with legitimizing broad delegations throughout its history. 190 Over the years, reasoned decision making emerged as an important means of *52 limiting discretion and improving the quality of agency
policymaking by requiring agencies to justify their actions.191 And while these reason-giving requirements exist in part to facilitate judicial review of agency action, they also have the intrinsic value of promoting agency experts' exercise of their discretion in a thoughtful, principled fashion.192 These requirements come from two sources: the Administrative Procedure Act193 supplemented by the requirement from SEC v. Chenery Corp. (Chenery II)194 that agency actions are valid only if they can be upheld according to the rationale given by the agency at the time the decision was made.195

*53 The APA procedures for informal rulemaking employ two strategies for ensuring that agencies can engage in a process of reasoned decision making.196 First, the APA has a notice requirement, which is designed to broaden the range of information and perspectives the agency must take into account197 in order to promote more informed decision making.198 To ensure these goals are met, the APA demands that an agency's notice must “fairly apprise interested persons of the subjects and issues” at stake whenever they intend to engage in a rulemaking199 and indicate the rulemaking's legal and factual basis as well as its policy purpose.200 Thus, the APA aims to ensure that agencies have before them all relevant information when making policy decisions.

The second element of the APA's strategy for channeling discretion is the obligation that agencies issue a public statement when announcing a final rule.201 Just as written judicial opinions demonstrate that a court's decision is supported by facts, law, and *54 precedent, these statements of purpose serve to demonstrate that agency officials considered all of the information before them and engaged in a “process of reasoned decision-making.”202 To that end, the statement must include the rule's basis and purpose as well as a justification of the decisions that led to its adoption, and it must “indicate the major issues of policy that were raised in the proceedings and explain why the agency decided to respond to these issues as it did.”203 Moreover, under Chenery II, only if the purpose provided by the agency constitutes a valid justification for the decision will it be legitimate.204

Rules should similarly dictate that the Attorney General or FBI Director provide notice explaining the reasons for any proposed changes to the Guidelines or the DIOG and a justification for the ultimate decisions that demonstrates that all of the relevant available information was taken into consideration and that there was a valid basis for the change. The distinctions between traditional administrative law and national security administration require, however, some adjustments to the usual procedural design. While most agency notices of proposed rulemaking are part of the public record and freely available for wide dissemination, the classified nature of much of the FBI's activity and some of the rules contained in the DIOG requires that the dissemination of the notice and justification will often be limited to individuals with the necessary security clearance.205 Imagine, for example, a proposed rule-change designed to modify surveillance operations to make them less likely to be detected by the target. Publicizing that intention and *55 the resulting rule might undermine entirely the purpose of the revision.

While the effect of limited dissemination of notice and justification will not be as robust as a process that is entirely transparent, limits on dissemination need not render written notice and justification worthless. In addressing concerns over excessive discretion, the crucial elements of the relevant administrative strategy are (1) that the notice broaden the range of information and perspectives that the agency considers and (2) that the written justification demonstrates that the agency's decision enjoys sufficient factual and legal support. To accomplish this, the notice and justification must go to individuals or entities whose participation would serve to expand the information and perspectives available to the decision makers and whose scrutiny of the ultimate justification would encourage the adoption of rules supported by reasoned argument and available evidence.
Candidates to receive this notice and justification are both inside and outside the Justice Department. The Justice Department’s Civil Rights Division and Office of Privacy and Civil Liberties as well as the National Security Division could be invited to comment. Similarly, other members of the intelligence community, such as the Office of the Director of National Intelligence should be involved. But the Privacy and Civil Liberties Oversight Board (PCLOB) should contribute its perspective as well. This is not meant to be an exhaustive list of possibilities. Nor will each of these institutions necessarily take on the role of civil-liberties champion or do so effectively. For now, it is enough to say that so long as they have access to the relevant information and bring a perspective different from the one within the Attorney General’s office or the FBI, their participation would help to channel discretion in a productive direction.

The 2008 modifications to the Guidelines provide a concrete example of how the requirement that the Attorney General provide reasons justifying amendments could impact the process. One proffered justification for amending the Guidelines in 2008 was that they were necessary to provide tools the FBI needed to support its preventive role. But the then-existing Guidelines already described prevention as the FBI’s “central mission,” and included several provisions added in 2002 that were aimed at empowering and enabling this aspect of the FBI’s activities. In asserting that the 2008 Guidelines were necessary for terrorism prevention, the Bureau had no obligation to explain how its activities were unacceptably constrained by the rules that were then in effect, which were drafted for the same purpose. Under these proposed rules, the Attorney General would have had to make that case, allowing those entities that were notified the opportunity to question the need for changes and possibly even make those concerns public. And if the Attorney General could not do so satisfactorily, it would have made altering the Guidelines a much more controversial proposition.

Requiring written justification of the final rules also might have had an impact on the 2008 revisions. On the one hand, the Justice Department downplayed the extent to which the Guidelines expanded the FBI’s powers. The new Guidelines were characterized as merely consolidating several existing sets of rules without making substantive changes. In fact, Justice Department officials asserted that assessments were nothing new, and that under the pre-2008 Guidelines it could already conduct assessments using pretext interviews, physical surveillance, and the tasking of informants. But there was no “assessment” level in the pre-2008 Guidelines. Indeed, they explicitly prohibited both pretext interviews and physical surveillance until a preliminary inquiry-a predicated investigation-had been opened. “Threat assessments” were permitted in some contexts under the 2003 National Security Investigation Guidelines, but by the FBI’s own admission, some of the techniques available in today’s assessments were prohibited in that context. Had the Attorney General or FBI Director been required to explain exactly what changes were being made and provide the rationale for the new rules, the creation of assessments in their current form-the central innovation of the 2008 Guidelines-might have met with more resistance. In particular, the ways in which assessments expanded the FBI’s powers, for example permitting use of several investigative techniques historically reserved for predicated investigations, would have become clear and consequently would have been subject to closer examination. As it was, the Attorney General could simply disclaim the idea that new powers were being granted without having to substantiate that statement.

In addition to any substantive differences to the 2008 Guidelines that would have resulted, a notice and justification requirement would have provided other concrete benefits. First, it would have guaranteed that decision makers had the benefit of additional perspectives while they were still developing the policy. It is much easier—and therefore more likely—for policymakers to incorporate alternative perspectives into policy still being developed than once that policy is nearing its final form. Second, obligating the Attorney General to address the information that was submitted to him and to explain in writing why the Guidelines should be implemented in his chosen format would force him to digest that information and therefore might actually result in a more informed decision. Third, it would have added additional
legitimacy to the final product if the 2008 Guidelines development had been based on specific, reasoned arguments regarding the need for modifications.

C. Participatory Policymaking

Nowhere does the Constitution provide for the existence of administrative agencies, much less for specific means of ensuring that their actions do not infringe on fundamental rights or that they are subject to democratic accountability. The administrative state has grappled with this “democracy deficit” almost since its inception. One means employed to address it has been through increasing opportunities for broad participation in agency decision making. Designing ways to employ these strategies for increased participation could lend increased democratic legitimacy to the intelligence-collection realm.

1. Administrative Agencies' Democracy Deficit

Administrative agencies initially were viewed either as entities merely implementing congressional will or as bastions of expertise, making decisions on the basis of scientific or technical knowledge. By the latter half of the twentieth century, however, it had become clear that many congressional delegations are vague and that many agency decisions cannot be resolved definitively through substantive expertise. Instead, such decisions often rest on subjective judgments about policy priorities, the value of human well-being, and who should bear the costs of inevitable risks. The Guidelines are thus not alone in their undemocratic nature. This democracy deficit “has spawned an extensive literature concerning the legitimacy of the administrative state.” Indeed, Professor Jody Freeman has suggested that “administrative law scholarship has organized itself largely around the need to defend the administrative state against accusations of illegitimacy” based on the unaccountability of agency officials, a lack of transparency, and limited opportunities for public participation.

*60 2. Increasing Democracy, Increasing Participation

One answer to the democracy deficit implemented by the administrative state has been to boost the democratic pedigree of agency rules by insisting on broad participation. A variety of procedural and doctrinal rules in the administrative state promote the broad participation of interested stakeholders. As with limits on discretion, the informal rulemaking process set out in the APA promotes participation through the implementation of strategies that can help improve the democratic pedigree of the Guidelines and the DIOG.

The first element of the strategy for improved democratic legitimacy is that the notice of proposed rulemaking itself must be provided in such a way as to facilitate meaningful participation, such as including the legal and factual basis for the proposed rule and the data on which the agency relied in making its proposal. These mandates ensure that stakeholders who want to participate have enough information to permit them to raise objections, provide additional information, or offer alternative perspectives. The notice, therefore, not only alerts diverse interested parties that there is a decision being contemplated for which they might want to provide input, but also ensures that input can be meaningful. Second, the agency must actually consider this input. A final rule's statement of basis and purpose must “indicate the major issues of policy that were raised in the proceedings and explain why the agency decided to
respond to these issues as it did.” 230 Thus any failure to take into account relevant comments can invalidate the rule. This risk of invalidation discourages the development of rules that do not take all relevant perspectives into consideration. 231

Again, these requirements will also result in rules that enjoy more democratic legitimacy than a rule prepared without such input. Affected parties are more likely to view agency decisions as legitimate if the process provides for a meaningful opportunity for presentation and consideration of their views. And if the rules are considered legitimate, the FBI will be much more likely to enjoy the full support and cooperation of the communities it is policing, leading to more effective intelligence collection.

Devising rulemaking mechanisms that are inclusive and allow for meaningful input from interested stakeholders presents a challenge when it comes to the domestic-intelligence regime because secrecy presents a formidable barrier to inclusion. 232 Even with respect to rules that are themselves public, such as the Guidelines, robust public participation in the process is impractical because there is insufficient public information about how those rules are implemented. The public may know, for example, that FBI agents are permitted to attend any religious service that is open to the public. 233 But it will not know how *62 often agents engage in this activity, what information they collect, or what is done with the information. Rules that are themselves secret, such as portions of the DIOG, present an even more formidable challenge. To be sure, a strong case can be made that the existing levels of secrecy—with respect to both the rules and the policy implementation—are excessive. 234 Unless and until that secrecy is reduced, however, intelligence-collection policies will struggle to gain the benefits that inhere in broadly inclusive agency rulemaking.

Yet, the Justice Department clearly hopes to realize at least some of the benefits of democratic input when it comes to the Guidelines. Recall the Department’s 2008 briefing of relevant congressional committees and other interested parties prior to the Guidelines’ adoption. 235 The Bureau did the same thing in 2001 when it revised the DIOG. 236 These meetings allowed the FBI to characterize these documents as rules developed with input from an array of stakeholders and thus deserving of the enhanced legitimacy that broad participation confers. 237 So even if the consultations themselves failed to result in meaningful participation beyond the Justice Department, 238 they indicate recognition that agency decision makers desire (at least the appearance of) an inclusive process.

The benefits of participatory decision making require a more robust process than the one undertaken in 2008. Recognizing that *63 the public at large will not be able to play a meaningful role, a second-best measure is to seek out proxies for points of view currently not formally represented in the process of developing the Guidelines and the DIOG. An example comes from Professors DeShazo and Freeman’s empirical case study of the licensing practices of the Federal Energy Regulatory Commission. 239 In the amended Electric Consumers Protection Act of 1986, 240 “Congress [required the Commission] to consult with fish and wildlife agencies prior to issuing licenses [and demanded] that nondevelopmental values be given ‘equal consideration’ with power concerns.” 241 DeShazo and Freeman explain that this amendment was specifically intended to strengthen the role of resource agencies in the Federal Energy Regulatory Commission’s decision-making process 242 and that it did in fact have the desired effect. 243

*64 A procedural regime governing the Guidelines’ development could similarly mandate the participation of particular entities that will bring alternative perspectives to the discussion. Each of the entities noted as a potential recipient of notice and source of information could play this role. One especially promising candidate for this role is the Privacy and Civil Liberties Oversight Board (PCLOB), a statutorily created independent agency charged with ensuring that privacy and civil liberties concerns are considered in the development and implementation of laws, regulations, and policies related to terrorism. 244 The PCLOB took several years to get off the ground. 245 In the wake of recent revelations regarding NSA
surveillance, however, it has demonstrated its ability to participate in the surveillance-policy conversation by insisting on a classified briefing about the controversial surveillance programs, meeting with the President, holding a public hearing seeking concrete suggestions for improving the civil liberties protections included as part of those programs. Based in part on what members of the Board learned at that meeting, the PCLOB issued a detailed report recommending specific changes to one existing surveillance program and anticipates issuing similar reports about other programs. Whether any of the PCLOB's recommendations will be adopted remains to be seen. But giving the Board an official role in the process in which the Guidelines and the DIOG are formulated would ensure that the civil liberties point of view is represented.

This participation could take one of several forms, ranging from an opportunity to express views to veto power. If the requirement is modeled on the concept behind the participation requirements included in notice-and-comment rulemaking, then it will end up somewhere between these two extremes, with something like the following arrangement: The PCLOB would have the opportunity to submit its perspective; having received this input, the Attorney General or the FBI Director would be required to demonstrate that it was taken into consideration.

Finding ways to broaden the perspectives involved when it comes to the DIOG is particularly important. Currently, nobody outside the FBI must be involved. Ensuring that alternative perspectives are voiced and requiring that the final rules reflect, or explain why they fail to reflect, these perspectives may provide some of the benefits of the multilateral, deliberative process that truly pluralist rulemaking procedures promote.

*66 D. Reconciling Conflicting Missions

Administrative governance also has developed ways to address the challenges posed when agencies are responsible for pursuing multiple, competing goals. The FBI must carry out its mission of preventing security threats from manifesting while simultaneously protecting fundamental rights. The principles behind the administrative state's tactics for reconciling conflicting missions offer ideas about how to implement structural checks to prevent the FBI's intelligence-collection mission from overwhelming these other important interests.

1. Juggling Mandates

Conflict among agency missions comes about when one or more statutes issue mandates to a single agency that come into tension with one another or when government-wide mandates conflict with the primary goals of individual agencies subject to those mandates. In one example, the National Park Service must protect the natural resources of the parks while simultaneously developing facilities for visitors. Similarly, the Fish and Wildlife Service is required to manage wildlife refuges for the conservation of plants and animals while also providing for recreation on those refuges. And the National Environmental Policy Act of 1969 (NEPA) requires “all federal agencies” to minimize the environmental impacts of their actions. For an agency focused on, for example, building roads through environmentally sensitive territory, the charge to protect the environment can be at odds with this focus.

The FBI's mandate to protect civil liberties can be viewed as a “secondary” mission-one that frequently comes into tension with its primary mission of preventing security threats. Studies show that an agency will focus on what it considers to be its primary mission, and it will shirk on performing “secondary” or less easily evaluated goals. As a secondary mission, protection of civil liberties is, therefore, sure to be short- changed in favor of security in the same
way that environmental concerns have so often gone under-addressed in favor of development or other economically profitable activities.

2. Relieving the Tension Among Multiple Missions

Fortunately, several administrative law strategies suggest ways to ensure that the Guidelines regime sufficiently takes into account civil liberties concerns as well as security concerns. Though all of the options discussed below are possible paths to follow, the final two approaches discussed below seem particularly promising.

Congress Reclaims Authority. One option, of course, is for Congress simply to relieve an agency of responsibility for one of the competing goals, reclaiming that decision-making authority for itself. Following revelations of civil liberties violations in the 1970s, Congress reclaimed some decision-making authority regarding the executive's surveillance powers by enacting the Foreign Intelligence Surveillance Act (FISA). Or Congress could generate more piecemeal limitations, barring particular techniques that pose threats to civil liberties, or defining the circumstances under which such techniques could be used. Congress could, for example, statutorily reinstate the rule regarding the use of undercover agents to investigate First Amendment protected activities as it existed in the Guidelines in 2001, which required that the FBI have probable cause or a reason to believe a crime had been committed before sending an agent into the meetings of a religious or political group.

Congress need not legislate to bring such changes about. If Congress wanted to alter particular investigative tactics, or even to pressure the Justice Department to adopt of its own volition the type of procedural framework suggested in this Article, it has an array of tools at its disposal to press for its desired policy change. Just the threat of legislation, so long as it is credible, can spur executive action. Recall that the original Attorney General's Guidelines were implemented to sap the momentum from Congress's efforts to enact a legislative charter for the FBI. So long as the option of enacting an FBI charter remains a viable means for Congress to limit the Attorney General's discretion when it comes to FBI investigations, the threat of such legislation can be used to press for Congress's desired policy outcomes. Congress possesses carrots as well as sticks-its control over the FBI and Justice Department's budget also can impose a great deal of pressure for policy change. Given the political economy of this policy area, however, reliance on Congress to reconcile the tension between the FBI's security mission and civil liberties is not the most promising route.

Separate Agency Functions. Another way that the administrative state deals with competing mandates is to separate agency functions, assigning one mandate to another (new or pre-existing) agency and leaving each free to focus solely on its own particular mandate. The APA's requirement that investigative and adjudicative functions be separated from one another, thereby insulating some decision making from possibly biased influences, is a way to implement this division-of-functions idea within a single agency. Professor Rachel Barkow has advocated, for example, for the separation of adjudicative and enforcement functions within prosecutors' offices. And in the domestic investigative context, the United Kingdom offers an illustration. Rather than relying on one agency both to enforce criminal laws and to collect intelligence, those functions are divided between two different agencies. The police forces investigate crimes and enforce criminal law, and MI5 collects intelligence. Some commentators have argued that the United States should consider more closely the idea of spinning off the FBI's intelligence-collection function into an independent agency. This alone would not, of course, address many of the concerns that the FBI's current powers raise. But it is possible that, recognizing the special threats to civil liberties that intelligence collection poses, an agency designed solely for that purpose would be subjected to more stringent limits. Indeed, to prevent overreaching, MI5's expansive intelligence-
collection powers do not include arrest or detention authority. Thus powers that are necessary for successful anti-crime efforts could nonetheless be off-limits or curtailed for the intelligence agency.

This division-of-functions solution, whether within or between agencies, is also unlikely to garner much support in the Guidelines context. As an initial matter, congressional passivity with respect to intelligence oversight will undermine any legislative efforts in this direction. But more importantly, many of the reforms to the intelligence community's structure in the past decade-plus have been explicitly designed to consolidate, rather than separate, functions. Perceived information-sharing failures prior to 9/11 led to a chorus of calls for breaking down barriers both within and between agencies, and both Congress and the executive branch have responded. The USA PATRIOT Act's removal of the so-called “wall,” which barred coordination between law enforcement and intelligence officials, is perhaps the most well-known, though by no means the only, post-9/11 change along these lines. Regardless of the salutary impact that separation of functions might have on civil liberties, the perceived security value of consolidation means that neither Congress nor the executive seems likely to reverse this trend.

Generating Information. More promising models of reconciling conflicting priorities are focused on agency culture, rather than agency structure. One mechanism for placing pressure on agency culture and prompting decision makers to consider factors that they otherwise might not give much weight is a requirement that an agency generate certain types of information. According to Professor Eric Biber, for example, “[a] major goal of NEPA was to force agencies that formerly had focused too heavily on primary missions such as highway construction, water-project development, or the extraction of natural resources, to also consider the impacts of their actions on the environment.” To accomplish this goal, NEPA requires all federal agencies proposing actions that will “significantly [affect] the quality of the human environment” to prepare an Environmental Impact Statement and make copies available to the public for written comments. These statements augment the information available to agencies, including the possible impacts on the environment, and proposals about how to avoid adverse environmental effects. And commentators agree that NEPA has been successful in integrating environmental goals into agency decision making. Similarly, all agencies—including the FBI—must generate a Privacy Impact Assessment (PIA) for “any substantially revised or new Information Technology System” that collects, maintains, or disseminates personally identifiable information from or about members of the public. And the Department of Homeland Security's Office of Civil Rights and Civil Liberties generates Civil Rights and Civil Liberties Impact Assessments when required to do so by statute, when they are requested by Department officials, or when the Officer for Civil Rights and Civil Liberties believes it appropriate.

Generating these assessments not only facilitates oversight efforts from the public, the legislature, or internal watchdogs, but-like requiring written justifications for changes-it has other benefits as well. As an initial matter, it forces agency decision makers to consciously consider the impact their proposed policy will have. As one set of commentators put it, “a systematic review of potential impacts during the planning process can focus the attention of decision makers on issues that they would otherwise deem to be outside their agency's mandate.” Requiring that effort will, at times, lead to agency choices more solicitous of the issue on which the assessment is focused. Decision-makers might simply need to be made aware of the impact of their choices. In addition, they will recognize that the substance of the assessment will be subject to scrutiny and, perhaps, criticism that they would rather avoid. And by ensuring that this information is before the decision makers while they are engaged in the decision-making process—rather than after the fact—makes the exercise all the more likely to have an impact. In addition, DHS's Privacy Office Official Guidance on Privacy Impact Assessments notes that the use of PIAs “demonstrates to the public and to Congress” that the new systems “have consciously incorporated privacy protections,” contributing to the legitimacy of the systems.
In order to ensure that the Attorney General or FBI Director consider explicitly specific “secondary” goals, he or she should be required to prepare a “Civil Liberties Impact Statement,”[73] articulating the likely effects of any proposed changes to the Guidelines. Requiring the Attorney General to consider, and to explain, whether the cost to civil liberties of any particular rule or tactic outweighs its investigative benefits is sure to raise the profile of civil liberties protection in the decision-making process. And while these Statements will not include the detailed scientific analysis that forms part of Environmental Impact Statements, they will identify the potential civil liberties impacts of proposed rules and force government officials both to note those impacts, and to think about what steps can be taken to mitigate them.

Inter-Agency Lobbying. A final mechanism the administrative state has used successfully to force agencies to consider specific, under-emphasized perspectives is for the political branches to enlist other agencies to police the primary decision-making agency. This idea, too, has potential in the Guidelines context. In some ways, this approach is simply a form of expanding the scope of participation, including a “lobbying” agency in the decision-making process to represent a particular interest that the decision-making agency is required to consider. [283] In other words, for the lobbying agency, its primary mission is to promote an interest that may be a “secondary” goal to the decision-making agency. This approach can differ from merely expanding participation in the process in that it envisions a more active role for the lobbying agency than merely providing a particular view to the decision maker. Again, a recommendation already mentioned provides an example. Recall Professors DeShazo and Freeman’s study about the licensing practices of the Federal Energy Regulatory Commission and the impact of Congress’s requirement that the Commission consult with fish and wildlife agencies prior to issuing licenses. [284] They concluded that this requirement had a real impact on the Commission’s treatment of the fish and wildlife agencies’ concerns. [285]

*74 Conferring a role in the decision-making process on an agency whose priority is the protection of fundamental rights might mitigate concerns that the FBI’s primary mission will unnecessarily endanger civil liberties. The Attorney General might, for example, be required to include officials from such an agency in the process of devising the FBI’s investigative rules-to give them a seat at the table. Just as including the Fish and Wildlife Service in agency decision making ensures that decisions take animal habitats into account, including an agency like the PCLOB could play a similar role with civil liberties concerns in the intelligence-collection context. [286] Members and staff of the board-many of whom, unlike staff at the Office of the Director for National Intelligence (ODNI) or FBI charged with protecting civil liberties offices, are drawn from the privacy and civil liberties advocacy community [287]-could raise civil liberties concerns that particular rules present, offer alternative means of achieving the FBI’s desired ends, suggest procedural protections that should accompany particular rules, or argue that certain rules should not be approved at all. Most importantly, ensuring an entity such as the PCLOB a seat at the table means that there is a voice actively involved in the process whose primary concern is not necessarily the prevention of terrorist acts.

While this type of interagency influence exertion can happen informally, [288] the regulatory or legislative creation of more hierarchical forms of agency interaction to vindicate “secondary” goals is likely more effective. Such a hierarchical structure causes decision makers to regulate in the “shadow” of that lobbying agency, prompting the decision-making agency “to internalize the *75 secondary mandates.” [289] Consider the role of the OIRA, through which the Executive Office of the President monitors regulation. [290] A presidential order places OIRA in a hierarchical position over federal agencies to evaluate whether the benefits of agencies’ proposed rules exceed their costs. [291] And OIRA has the power not just to make suggestions for modifications but actually to block implementation of an agency regulation on this basis. [292] This “veto” power requires those agencies to take into account what they might consider a secondary goal-efficiency-when contemplating regulatory action. [293]
As with the requirement that the Attorney General or FBI Director consider input from particular entities, the impact and effectiveness of this model would be highly contingent on the degree to which Justice Department officials were obligated to take the PCLOB's opinions into account. An agency statutorily empowered to overrule DOJ proposals would have enormous practical effect. But such drastic (and implausible) measures are not required. Again, there are a range of possible roles for the PCLOB. For example, it could simply be given a seat at the table during the formulation of the Guidelines, allowing its representative to raise civil liberties concerns. Or a stronger thumb on the scale of civil liberties might be to require a report to Congress about any instances in which the PCLOB and the Attorney General or FBI Director are unable to reach agreement on a particular issue.

In sum, the procedural framework for the Guidelines should explicitly require that the Attorney General or FBI Director take into account the civil liberties costs when weighing policy options by preparing a “Civil Liberties Impact Statement” detailing the likely impact of the proposed changes on fundamental rights, and should empower the PCLOB to play an active role in formulating the Guidelines. These suggested procedures will not eliminate all concerns about civil liberties raised by the Guidelines and the DIOG. But in a context where preferred methods of rights protection break down, they offer a second-best option.

IV. Theoretical and Practical Objections

Even if one concedes the potential value of the above reform suggestions, a couple of questions might arise. First, if the answer to the governance gap in intelligence collection is to import administrative law principles, why are existing proposals in that vein insufficient? And second, in the absence of judicial review and public scrutiny facilitated by transparency, how will the procedural requirements suggested here be enforced? This Part will address each of these questions in turn, arguing that the reforms suggested above provide a better means of addressing civil liberties concerns than existing reform proposals, and that there are available mechanisms that can enforce compliance with the recommended governance regime.

A. Distinguishing Alternative Proposals

Recognizing the inadequacy of the current regime, some scholars have suggested that terrorism be treated as an administrative problem. This is a profound insight. After all, the Justice Department and the FBI are agencies, and there is an entire body of law whose raison d'etre is to improve agency governance regimes and ensure reasoned exercise of broad delegations of power. But existing reform proposals do not confront the Guidelines' governance challenges that implicate civil liberties concerns. Some call for the development of mechanisms to improve our ability to assess the actual risk posed by terrorism. Others argue that responses to terrorism should take into account the unusually high psychological costs of terrorism. Yet another approach looks for governance solutions through the structure of congressional oversight of intelligence operations.

While some of these suggestions may prove beneficial to counterterrorism efforts more generally, none offers solutions to the civil liberties threats posed by the Guidelines and the DIOG. Indeed, they are not focused on meeting those challenges. Rather, they are concerned with broad-gauge adjustments to the government's general approach to counterterrorism. As a result, they do not purport to offer means of channeling agency discretion, enhancing the Guidelines' democratic bona fides, or ensuring that concerns about fundamental rights are given the attention they deserve.
One scholar does consider in detail what a regulatory approach to intelligence collection might look like on the ground. Professor Sam Rascoff advocates importing into the intelligence-collection context several of the traditional means of constraining agency action, such as centralized cost-benefit review of agency rules to ensure that rules are “rational”- which he defines as efficient, effective, and sufficiently rights-respecting - judicial review, and increased public participation and transparency in intelligence agency decision making. These mechanisms, he asserts, will create a “risk-management approach to counterterrorism” akin to other areas of regulatory endeavor. Rascoff predicts that this approach will increase the accuracy, efficiency, and usefulness of the intelligence that is collected, and it may in fact do so.

In addition to yielding gains in the quality and efficiency of intelligence collection, Rascoff expects his approach also to promote the protection of rights and check abuse or illegality; it is here where his proposal, in my view, falls short. In fact, when it comes to addressing the civil liberties concerns inherent in the Guidelines regime, his approach contains two flaws. First, it is undertheorized. The traditional tools of administrative law-such as cost-benefit analysis, judicial review, and public participation-simply cannot function effectively in the intelligence-collection context. Their reliance on transparency, public scrutiny of agency action, and judicial review renders them largely inapplicable to intelligence collection. To be sure, Rascoff recognizes the distinctiveness of the context and points to ways to modify these tools to operate more effectively in the intelligence-collection context. Judicial review, for example, would be performed by the secret Foreign Intelligence Surveillance Court (FISC), rather than traditional federal courts; and cost-benefit analysis would be done by experts in the ODNI rather than OIRA. But these adjustments do not go far enough. Instead, the obstacles to regulating intelligence collection require the formulation of new institutional designs custom-tailored to meet the challenges that intelligence-collection regulation presents. Second, and perhaps more importantly, embedded in Rascoff’s proposal is an overly optimistic perspective on the premise-a premise belied by history-that the intelligence community can be relied upon to develop and enforce civil liberties protections on itself.

Consider first the idea of centralized review of intelligence policy in the ODNI to promote intelligence-collection rules that are “rational”-efficient, effective, and sufficiently rights respecting. Just as OIRA subjects agency regulations to cost-benefit analysis, ODNI would review intelligence policies for rationality. And because this rationality includes a mandate to seek the appropriate balance between security and liberty, the argument goes, it will ensure that privacy and liberty interests are not ignored.

Looking to ODNI review as a means of rationalizing intelligence policy, however, is not a solution to civil liberties concerns. Unlike regulatory action whose economic impact can be measured, the liberty costs of various security policy options are often either speculative or a matter of subjective valuation. How would ODNI weigh the privacy harms caused by surveillance? How many unnecessary additions to a terrorist watchlist are justified by the detection of an incipient plot? Moreover, even if such considerations could be quantified, ODNI will be unable to know whether different, less intrusive, policies could produce the same positive security result. Merely instructing ODNI to take liberties costs into account may result in more liberty-solicitous policy. But because these costs cannot be objectively evaluated and intelligence community members are likely to assign lower value to them than the population at large, cost-benefit review is not a reliable mechanism for protecting individual rights.

Moreover, ODNI itself is a member of the intelligence community. It may be the case that ODNI “enjoys sufficient distance from the various intelligence agencies” that it cannot be co-opted by any one element of the intelligence community. But ODNI need not be “captured” by the FBI to fall prey to the same pro-security biases. ODNI’s website
identifies its vision as, “A Nation made more secure because of a fully integrated Intelligence Community.” 316 Indeed, the staff of ODNI itself, at least initially, was drawn from other agencies within the *81 intelligence community. 317 This vantage point inside the national security apparatus means that ODNI's assessment of the value of proposed intelligence-collection policy and the costs of resulting privacy and liberty sacrifices will not differ markedly from that of the FBI itself. Because ODNI officials share the FBI's primary mission of ensuring security, it is unclear why they would value civil liberties any differently than FBI officials or the Attorney General. 318 And because quantifying the value of civil liberties is an inherently subjective exercise, moving responsibility for doing so from one element of the intelligence community to another seems unlikely to generate significant rights-protection improvements.

Even if the ODNI is able to set aside its security hat for the purpose of evaluating the civil liberties implications of particular policies, it will receive a distorted view of the policies in question. ODNI will not have the benefit of information provided by stakeholders outside the intelligence community. Intelligence community officials are likely-either consciously or unconsciously-to overvalue their own role and to present information in ways that support the intelligence-collection policies they favor. Without subjecting such information to scrutiny outside the intelligence community or considering outside views, any relevant flaws or biases in the information presented to the ODNI will remain unchallenged. So while centralized review would impose an additional layer of bureaucratic scrutiny of the Guidelines, the additional review would not replicate the benefits that centralized review has produced in other parts of the administrative state, at least when it comes to civil liberties, because the ODNI cannot offer a neutral, dispassionate evaluation of the benefits and drawbacks of intelligence-collection policy.

*82 The suggestion that the FISC approximate the role of traditional judicial review of agency decision making to impose constraints on discretion will also fail to result in the preservation of civil liberties. As an initial matter, it is unclear what the extent of the FISC's review might be. Traditional judicial review of administrative rules asks whether an agency's action is consistent with the Constitution and its statutory mandate or whether it is arbitrary or capricious. 319 But when it comes to most intelligence-collection rules, there is no constitutional or statutory standard against which a court could measure agency compliance. 320 One proposed solution to this baseline problem is to have the FISC review policy for whether it is consistent with the intelligence agencies' own stated objectives. 321 Again, this proposal fails to account for the fact that when the intelligence community is left to determine the rules of its own conduct, concerns other than security will get short shrift. By asking intelligence agencies to identify their own objectives and then subjecting their efforts to meet those objectives to judicial review would replicate the current situation—where the constraints on agencies are limited to those that they agree to place on themselves—but with the added legitimating feature of judicial imprimatur.

Another barrier to enlisting the FISC in intelligence-collection governance is that the intelligence-collection activities governed by the Guidelines extend beyond the scope of the FISC's jurisdiction. The FISC oversees electronic foreign intelligence surveillance and physical searches of premises connected with foreign powers. 322 It has no role in overseeing purely domestic *83 surveillance of Americans absent probable cause that those Americans are agents of a foreign power. 323 The content of the Guidelines and the activities they regulate—such as physical surveillance of Americans, infiltration of religious or political groups, the use of informants, requests for internet history—rarely fall within the FISC's jurisdiction. Individuals who wish to challenge FBI activity—if they can establish standing—do not have access to the FISC. 324 Thus, it is unclear what role the FISC could play in reviewing many activities in which the FBI engages.

The FISC, too, is likely to share the FBI and ODNI's bias toward the security mission. Unless a recipient of a FISC order challenges the legitimacy of that order, proceedings in the FISC are not subject to an adversarial process. 325 Instead, like magistrate judges considering whether to issue traditional search warrants, FISC judges review unopposed government
applications for surveillance orders. The FISC thus receives only the Justice Department's perspective—heavily informed by the FBI's perspective—about any given rule. This concern is compounded by the fact that even the judges themselves largely hail from the law enforcement community—twelve of the fourteen judges who have served this year are former prosecutors and one is a former state police director. Moreover, once selected by the Chief Justice of the Supreme Court for FISC service, these judges are exposed to a constant stream of government applications to engage in foreign intelligence collection detailing just how dangerous the world can be and the important role that intelligence collection plays in combating those dangers. FISC involvement thus serves only to reinforce the pro-security perspective already embedded in the development of domestic-intelligence-collection policies.

Finally, Justice Department briefings to congressional committees and to interested nongovernmental organizations shortly before the current Guidelines were officially issued can be interpreted as evidence of “the possible emergence in the domestic intelligence arena of a new ethic of interest group representation.” This form of participation, however, left many interested parties unsatisfied and bore only “a passing resemblance” to traditional notice-and-comment rulemaking. These meetings permitted stakeholders to see (but not to copy or retain) a near-final draft of the document weeks before its implementation. But the Guidelines were entirely subject to the Attorney General's discretion with respect to whether to respond to or to take any views expressed at these meetings into consideration. This superficial involvement of interested parties does not provide sufficiently broad, meaningful participation outside the Justice Department and the FBI to alleviate the accountability and democracy deficit with the Guidelines and DIOG.

B. Oversight of Procedural Requirements

One possible objection to this Article’s proposals is that, in eschewing judicial review and conceding the secret nature of the Guidelines regime, they relinquish all means of enforcing their requirements. As with the proposed procedural rules themselves, however, the principles behind the administrative state's compliance mechanisms offer a (partial) solution.

Notice-and-comment rulemaking employs the transparency of the rulemaking process followed by public and judicial review to enforce the regulatory regime to which agency decision making is subjected. The public's role in that regime is to play watchdog. Because rulemaking and its results are conducted in a transparent fashion, interested stakeholders can be relied upon to object if they believe that an agency has not acted appropriately. The public will scrutinize not only proposed rules to ensure that they do not suffer from procedural, logical, or evidentiary deficiencies, but also any information that is made public as part of the process. If, for example, an Environmental Impact Statement predicts dire environmental consequences from a proposed agency action, environmental activists will use that Statement to lobby not only the agency but also Congress and the President to prevent the agency from taking the proposed action. Another important compliance-related element of public scrutiny is the fact that it can lead to legal challenges to agency rules. If a regulated entity believes that a regulation applied to it was adopted through flawed procedures, it can bring suit, thereby subjecting the regulation to judicial review. Through this review, courts serve to confirm that agency decisions are not unjustified exercises of discretion and that they followed the mandated procedures. Consequently, courts engage in a “searching and careful” review of the record an agency makes of its decision-making process and will invalidate the results of proceedings that are “arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law.” These standards are not necessarily particularly stringent, but like public scrutiny, they do ensure that agency decisions have been reached through the proper procedures and therefore that they are reasoned rather than arbitrary or irrational.
Perhaps just as importantly, the roles that the public and the courts play impose an important “prior restraint” on agencies. Knowing that their rules and the justifications that they offer for them will be public and potentially subject to judicial scrutiny, agencies will be more likely to be conscientious, hoping to ensure that their decision-making processes pass judicial muster. In other words, they will take any procedural requirements seriously from the outset knowing that, if they do not, any resulting rule could ultimately be invalidated.

Fashioning equally effective means of supervising Justice Department or FBI compliance with any relevant rules is a challenge, because the transparency that facilities both public scrutiny and judicial review is concededly difficult to replicate. Any proceedings regarding the Guidelines that take place outside of public view are shielded from the public-as-watchdog. If the relevant rules were legislatively mandated, challenges to any failure to abide by them theoretically could be reviewed by the courts. But even in the unlikely event that Congress imposes procedural requirements akin to those suggested here, such suits will fall prey to the same barriers that currently exist to challenging the FBI’s intelligence-collection activities—any individual or entity seeking to challenge the Guidelines or the DIOG on the grounds that they did not follow the required procedures would struggle to establish standing and to overcome the state secrets privilege.

As with the reforms suggested to channel the Justice Department’s discretion and to improve the participatory nature of the Guidelines’ development process, ensuring compliance with procedural requirements would necessitate a means of approximating the traditionally public and judicial roles. One option in this regard would be to enlist proxies within the executive branch to engage in scrutiny of the decision-making process, to inquire whether the process complied with any required procedures, and to consider whether the required statement(s) of justification are adequate. A government watchdog could be assigned to take the place of the public and judicial watchdogs that normally play this role. The Justice Department’s Inspector General (IG)—who is statutorily empowered to conduct audits, investigations, inspections, and reviews of Justice Department programs and to issue reports to Congress regarding the results of any investigations that it does conduct—might play a constructive role in holding the Attorney General and FBI Director accountable for following any applicable procedural rules. The IG investigates not only alleged violations of the law by DOJ employees, but also audits and inspects DOJ programs regularly. That office could perform reviews of the process employed each time the Guidelines or the DIOG are amended. Audits conducted by the IG would be especially effective in replicating the effects of traditional transparency if the results of those audits could be released publicly. Indeed, IGs have, at times, played quite important roles in uncovering violations of law and policy in pursuit of security. Perhaps more than any other oversight mechanism (with the exception of unlawful leaks of classified information), audit reports from the Justice Department’s Inspector General have shed light on the FBI’s investigative activities in the wake of 9/11. These reports, some of which revealed violations of law or policy, drew both public and congressional attention, and consequently prompted changes to internal FBI policy.

To be sure, even publicly released IG conclusions would lack some of the other compliance-enhancing characteristics of public and judicial scrutiny. The IG does not have the power to invalidate rules that are adopted through flawed procedures or lack sufficient justification. It can point out flaws and insufficiencies, but ultimately any findings or recommendations would be nonbinding. This absence of compulsory power sacrifices some of the sword-of-Damocles threat inherent in the promise of judicial review. If, however, the findings and recommendations can be made public, the threat of reputational costs to the FBI still imposes some ex ante incentive to comply with required procedures. And IGs have been particularly successful in generating public reports for reviews of even the most sensitive programs.

An alternative, though less promising, option would be to rely upon Congress to monitor compliance through either the General Accounting Office, which regularly audits and reviews agency programs on behalf of Congress, or the
congressional oversight committees themselves. This option is less promising than the IG, however, because Congress already has the power to insist on these types of procedures but has chosen not to intervene when it comes to the Guidelines and the DIOG. Indeed, in 2008, several members of the Senate Judiciary Committee voiced concerns about imminent changes in the Guidelines. The Attorney General was under no obligation to take those concerns into account, however, and he did not respond to the Senators' correspondence.

V. Conclusion

Domestic intelligence collection presents a challenge in a democracy. While it can play a crucial role in keeping our nation secure, it also poses threats to the very freedoms that make that nation worth defending. When the prevention of terrorism is viewed as a regulatory problem—one to be managed rather than defeated—the challenge becomes more manageable. Examining regulatory strategies developed over the past half-century in the administrative state provides a roadmap for the development of structural and procedural mechanisms to channel executive discretion into reasoned, evidence-based decisions; to include viewpoints from outside the intelligence community in the process; and to ensure that the Justice Department explicitly takes into account the civil liberties perspective. The need to substitute alternative mechanisms for tools that are often effective in governing agency action, such as judicial review and public scrutiny, means that they may be less directly effective in achieving their goals. But developing such a framework is a viable second-best option in a context where the traditional means of government oversight break down. It will contribute to generating the appropriate FBI for the twenty-first century, one that takes into account not only the nature of the disease, but also the potential costs of the cure.

Footnotes

a1 Visiting Assistant Professor of Law, Brooklyn Law School, J.D., NYU 2005, LL.M. NYU 2011. Thanks go to the Brooklyn Law School Junior Faculty Workshop and Visiting Assistant Professor Workshop, Amna Akbar, Miriam Baer, Rachel Barkow, Robin Effron, Barry Friedman, Liza Goitein, Aziz Huq, Sarah Light, Mike Livermore, Teddy Rave, Margo Schlanger, Shirin Sinnar, Mila Sohoni, Nelson Tebbe, and Adam Zimmerman for helpful comments and discussions, and to the Brennan Center for Justice and Brooklyn Law School for generous research funding.


2 See infra Part II (discussing the nature of intelligence collection).

3 See infra Part II.A (describing sophisticated data-mining tools that allow the government to “render a detailed dossier on any American”).

4 Infra note 63 and accompanying text.

5 See infra Part II.A (discussing the FBI's broad intelligence-collection powers and the negative impact those powers can have on individual liberties).

to previous movements for more health and environmental legislation); Cass R. Sunstein, Terrorism and Probability Neglect, 26 J. Risk & Uncertainty 121, 129-31 (2003) (describing the effect public fear and probability neglect play in movements for environmental and health legislation and speculating what that effect should mean for counterterrorism regulation); James B. Comey, Intelligence Under the Law, 10 Green Bag 2d 439, 442-44 (2007) (discussing the importance of adhering to legal institutions, even in the intelligence community).

See Shirin Sinnar, Protecting Rights from Within? Inspectors General and National Security Oversight, 65 Stan. L. Rev. 1027, 1027 (2013) [hereinafter Sinnar, Protecting Rights from Within] (discussing the legislature and courts' reluctance to restrict the executive's power when it infringes on civil liberties in furtherance of national security); Amna A. Akbar, Policing “Radicalization”, 3 U.C. Irvine L. Rev. (forthcoming 2014) (discussing the government's monitoring of religious groups in an effort to prevent members of those groups from turning to violence); Shirin Sinnar, Questioning Law Enforcement: The First Amendment and Counterterrorism Interviews, 77 Brook. L. Rev. 41, 42 (2011) [hereinafter Sinnar, Questioning Law Enforcement] (discussing the First Amendment concerns raised when law enforcement selects individuals for questioning based on their speech and associations); Rascoff, supra note 6, at 586 (stating that the protection of civil liberties should factor into the review of intelligence gathering actions).

See infra Parts II.A & II.B.1 (discussing the lack of judicial and legislative oversight over intelligence collection and the lack of practical restraints on the FBI when it prioritizes the collection of data over everything else).

See infra Part II.B.2 (describing the lack of judicial, political, and public oversight over intelligence collection).

See infra Part II.B.3 (describing the FBI's primary focus on national security and intelligence gathering, and the fact that privacy concerns are viewed as a hurdle to effective policymaking).


Id. Initially secret in its entirety, Freedom of Information Act requests prompted the Justice Department to release the DIOG—without significant redactions. See Muslim Advocates v. U.S. Dep't of Justice, 833 F. Supp. 2d 106, 109 (D.D.C. 2012) (concluding that the release of the DIOG under a FOIA request was proper, but that the FBI was permitted to redact certain portions of the DIOG); Elec. Frontier Found. v. U.S. Dep't of Justice, 826 F. Supp. 2d 157, 174 (D.D.C. 2011) (directing the Department of Justice to release more detailed descriptions of the U.S. and European Union’s discussions over the international exchange of personal information).

See infra Part II (discussing the original goal of the Levi Guidelines to “strictly curtail domestic intelligence investigations”).

A recent proposal to develop a risk-management approach to intelligence-collection governance by employing traditional agency oversight tools—centralized cost-benefit analysis, judicial review, and pluralist input into decision making, see Rascoff, supra note 6, at 633-47 (outlining a plan for regulatory governance of intelligence gathering agencies)—fails to incorporate effective civil liberties protections. See infra Part IV.A (discussing reform proposals that focus on improving the FBI’s ability to correctly assess the risks and account for the psychological costs of terrorism, while not addressing the civil liberties concerns the Guidelines implicate).

See infra Part II.B.3 (discussing the fact that the Guidelines and DIOG place civil liberties in a secondary role to the FBI’s intelligence gathering mission).


See infra Part III.A (describing the improvements that could be made by implementing administrative law strategies in the intelligence-collection context).
See infra Part III (describing the broad opportunities for participation in administrative rulemaking, and the democratic legitimacy that approximating such participation could create in the intelligence-collection realm).

See infra Part III.D (describing the methods that administrative governance uses to reconcile and balance competing policy goals).

See infra Part III.C.1 (describing the administrative state's response to the inherent democracy deficit present in administrative actions).

See infra Part III.C.2 (discussing methods that could be used to introduce administrative procedures into intelligence collection).

See infra Part III.D.2 (proposing solutions which would allow the FBI to balance competing intelligence collection and civil liberties interests).

See infra Part IV.A (describing the inadequacies of administrative law tools alone).

See infra Part IV.B (addressing the fact that by its nature intelligence collection cannot be as transparent as traditional administrative procedures, and suggesting alternative compliance methods which can still alleviate some of the governance concerns implicated by intelligence collection).

See infra Part II.A (discussing the nature of the FBI).

See DIOG, supra note 12, Preamble (discussing broadly the investigative role of the FBI).


The Church Committee exposed a litany of intelligence-collection programs (most implemented under infamous long-time FBI Director J. Edgar Hoover) in which the FBI used widespread surveillance to harass and discredit law-abiding—though often antiwar or civil rights-groups and individuals based on their political beliefs. See Supplementary Detailed Staff Reports of Intelligence Activities and The Rights of Americans: Book III: Final Report of the Select Comm. to Study Governmental Operations with Respect to Intelligence Activities, S. Rep No. 94-755, at 27 (1976) [hereinafter Church Comm. Report] (detailing the Committee's findings).


See AGG Compliance Report, supra note 30, at 59 (“Attorneys General and FBI leadership have . . . referred to the FBI's adherence to the Guidelines as the reason why the FBI should not be subjected to a general legislative charter or to statutory control.”).

For a detailed account of the historical evolution of the Guidelines, see Berman, supra note 32, at 8-25.


Michael Mukasey, U.S. Dep't of Justice, Attorney General's Guidelines for Domestic FBI Operations (2008) [hereinafter Mukasey Guidelines]. Attorney General Mukasey's Guidelines, which remain in operation today, consolidate the Guidelines for criminal and domestic intelligence investigations with those for National Security Investigations and Foreign Intelligence Collection Guidelines. Id. § I.D.1. They also incorporate rules formerly contained in the Supplemental Guidelines for Collection, Retention, and Dissemination of Foreign Intelligence; the Guidelines for Reporting and Use of Information Concerning Violation of Law and Authorization for Participating in Otherwise Illegal Activity in FBI Foreign Intelligence, Counterintelligence, or International Terrorism Intelligence Investigations; and the Guidelines for Reporting on Civil Disorders and Demonstrations Involving a Federal Interest. Id.

See Samuel J. Rascoff, Establishing Official Islam? The Law and Strategy of Counter-Radicalization, 64 Stan. L. Rev. 125, 162 (2012) (noting the friction with the Constitution's Establishment Clause caused by the government espousing particular interpretations of Islam). The Guidelines have been subject to much criticism along these lines since 9/11. See, e.g., Lininger, supra note 36, at 1231-54 (analyzing Attorney General Ashcroft's new Guidelines and the dangers they pose to religious freedom, the threat that they create for racial and religious profiling, and the risk of inefficient use of resources); see also Daniel J. Solove, The First Amendment as Criminal Procedure, 82 N.Y.U. L. Rev. 112, 151-59 (2007) (suggesting that First Amendment claims arising out of domestic investigations should be cognizable); Linda E. Fisher, Guilty by Expressive Association: Political Profiling, Surveillance and the Privacy of Groups, 46 Ariz. L. Rev. 621, 643-57 (2004) (arguing that the law should bar federal officials from engaging in surveillance triggered by First Amendment activity without reasonable suspicion of a crime).

See Rascoff, supra note 6, at 589 (describing the “doctrinal vacuum” in the law of domestic intelligence); Anjali Dalal, Administrative Constitutionalism and The Re-Entrenchment of Surveillance Culture 24-29 (Mar. 4, 2013) (unpublished manuscript), available at http:// papers.ssrn.com/sol3/papers.cfm?abstract_id=2236502 (arguing that contemporary constitutional norms surrounding surveillance practices are in part a result of the entrenchment over time of the policy preferences of the FBI, the Attorney General, and the President) (on file with the Washington and Lee Law Review).

See United States v. Miller, 425 U.S. 435, 442 (1976) (stating that disclosure to a third party negates any expectation of privacy, and therefore Fourth Amendment does not apply); Smith v. Maryland, 442 U.S. 735, 745 (1979) (stating that when Smith
dialed a phone number he was disclosing that number to the phone company, and therefore had no legitimate expectation of privacy over the fact that he dialed that number).

41 See, e.g., United States v. Skinner, 690 F.3d 772, 777 (6th Cir. 2012) (stating that DEA did not violate Fourth Amendment when using cell phone to track location without a warrant). The Supreme Court has not ruled on this question.

42 See Hoffa v. United States, 385 U.S. 293, 300-02 (1966) (stating that when an undercover informant was invited into the hotel room of the defendant, there was no expectation of privacy and therefore no Fourth Amendment issue).

43 See Church of the Lukumi Babalu Aye v. Hialeah, 508 U.S. 520, 531 (1993) (noting that laws of general application need not be motivated by an important government interest, even if they infringe on individuals' religious rights).

44 See, e.g., Holder v. Humanitarian Law Project, 561 U.S. 1, 130 S.Ct. 2705, 2730 (2010) (holding that the First Amendment burden imposed by the criminal material support statute was outweighed by the government's interest in national security). The Guidelines could be implemented in ways that are vulnerable to an as-applied challenge, but such challenges are likely futile for reasons explained in Part II.B.2.


46 Mukasey Guidelines, supra note 37, at Intro. B.

47 Id. § IV.

48 Id. § II.


50 Mukasey Guidelines, supra note 37, § II.

51 Id.

52 Id.


54 See Mukasey Guidelines, supra note 37, § II.A.4.e; J.M. Berger, Does the F.B.I. Have an Informant Problem?, Foreign Pol'y (Sept. 7, 2012), http://www.foreignpolicy.com/articles/2012/09/07/does_the_fbi_have_an_informant_problem (last visited Nov. 14, 2013) (discussing “growing media scrutiny” over the FBI's professional informants' conduct) (on file with the Washington and Lee Law Review). Historically, the FBI has infiltrated political and religious groups for the purpose of disrupting their operations and creating a chilling effect on antigovernment expression. See Lininger, supra note 36, at 1235-36 (discussing government infiltration of religious organizations); Frederick A. O. Schwarz Jr. & Aziz Z. Huq, Unchecked and Unbalanced: Presidential Power in a Time of Terror 31-36 (2007) (describing the Church Committee's findings regarding the
FBI's politically motivated investigations); Don Edwards, Reordering the Priorities of the FBI in Light of the End of the Cold War, 65 St. John's L. Rev. 59, 73 (1991) (describing politically motivated investigations in the 1980s).

See Mukasey Guidelines, supra note 37, § II.A.4.f (“Interview or request information from members of the public and private entities.”).

See id. § II.A.4.h (“Engage in observation or surveillance not requiring a court order.”).

Id. § II.A.4.c. As of 2009, the FBI planned to expand this data set to include tax records from nonprofit organizations, see U.S. Dep't of Justice, FBI, Nat'l Security Analysis Ctr., An Element of the FBI's National Security Branch (2006), and the Bureau has launched a database of biometric information, which includes “[d]igital images of faces, fingerprints and palm patterns.” Ellen Nakashima, FBI Prepares Vast Database of Biometrics: $1 Billion Project to Include Images of Iris and Faces, Wash. Post, Dec. 22, 2007, at A1.


See id. (describing the breadth of the FBI's data-mining project).

Id.

See Berman, supra note 32, at 13-21 (discussing the erosion of the protections provided by the Levi Guidelines).


See Greenwald, supra note 1 (discussing a court order that “compels Verizon to produce to the NSA electronic copies of ‘all call detail records or “telephony metadata” created by Verizon for communications between the United States and abroad’ or ‘wholly within the United States, including local telephone calls’” (internal citations omitted)).

See Solove, supra note 38, at 151-76 (discussing the First Amendment implications of modern intelligence gathering).
69 See Sinnar, Questioning Law Enforcement, supra note 7, at 69-71 (discussing the chilling effects on the Muslim community caused by law enforcement actions).


71 Id.

72 Reports suggest that the DIOG requires a Special Operations Review Committee (SORC) to approve surveillance inside mosques. See Jessica Chasmar, Mosques Off-Limits by Government Snooping Since 2011, IBD Editorial Says, Wash. Times (June 13, 2013), http://www.washingtontimes.com/news/2013/jun/13/mosques-limits-government-snooping-2011-ibd-editor/ (last visited Nov. 13, 2013) (discussing the intersection of FBI intelligence-gathering and religious observation) (on file with the Washington and Lee Law Review). But the rules regarding undisclosed participation in political and religious events are redacted from the public version of the DIOG pursuant to FOIA's exemption for law enforcement material, as are the provisions regarding the responsibilities and makeup of the SORC. See DIOG, supra note 12, §§ 16, 18.5.5.3, 18.6 (describing undisclosed participation procedures, the fact that the FBI should use the least intrusive method possible, and the authorized investigation methods used in preliminary investigations); 5 U.S.C. § 552(b)(7)(E) (2012) (exempting from disclosure information compiled for law enforcement purposes “if such disclosure could reasonably be expected to risk circumvention of the law”).

73 Lininger, supra note 36, at 1233-34 (citations omitted); Fisher, supra note 38, at 647-49 (“The chilling of protected expression that accompanies political surveillance impedes the group's ability to realize fully its political or religious purposes.”); see also Int'l Religious Freedom Report, Hearing Before the Subcomm. on Int'l Operations, House Int'l Relations Comm., 108th Cong. 67 (2002) (statement of Nihad Awad, Exec. Dir. of Council on American-Islamic Relations) (stating that at least three Muslim charities “have been effectively shut down”). Impediments to charitable giving interfere with Muslims' ability to practice the Islamic duty of zakat (alms giving). See id. at 67-68 (“These closures [of Muslim charities] have had a wide impact . . . . Donors view such organizations as essential to the ability of Muslims to practice the religious duty of zakat (alms giving), a pillar of their faith.”).

74 See Sinnar, Questioning Law Enforcement, supra note 7, at 69-71 (detailing evidence that members of the Muslim Community refrain from expressing their religious or cultural identities).

75 The Supreme Court has noted that protections against intrusive surveillance “become the more necessary when the targets of official surveillance may be those suspected of unorthodoxy in their political beliefs.” United States v. U.S. Dist. Ct., 407 U.S. 297, 314 (1972).

76 DIOG, supra note 12, § 4.3.3.2.4.


78 See DIOG, supra note 12, § 4.3.3.2.2 (“[I]f information about community demographics may be collected, it may be ‘mapped.’ Sophisticated computer geo-mapping technology visually depicts lawfully collected information and can assist in showing relationships among disparate data.”).

79 Id. § 4.3.3.2.1. This same idea was proposed by local law enforcement authorities in Los Angeles but ultimately abandoned when the Muslim and civil liberties communities noted that it was likely to alienate Muslim residents. See Richard Winton et al., LAPD to Build Data on Muslim Areas, L.A. Times (Nov. 9, 2007), http://articles.latimes.com/2007/nov/09/local/

80 DIOG, supra note 12, § 4.3.3.2.1.


82 Mukasey Guidelines, supra note 37, § I.C.3.


84 See, e.g., Mukasey Guidelines, supra note 37, § I.C.3 (“These Guidelines do not authorize investigating or collecting or maintaining information . . . solely for the purpose of monitoring activities protected by the First Amendment . . . .”); DIOG, supra note 12, § 4.2 (“[I]nvestigative activity may not be based solely on the exercise of rights guaranteed by the First Amendment . . . .”).

85 See Aziz Z. Huq, The Signaling Function of Religious Speech in Domestic Counterterrorism, 89 Tex. L. Rev. 833, 842 (2011) (“[L]aw enforcement and prosecutors turn to religious speech as a signal of terrorist risk.”). Note also that after the Supreme Court decision in Holder v. Humaitarian Law Project, any speech undertaken in coordination with a designated Foreign Terrorist Organization (FTO) is criminal. Holder v. Humaitarian Law Project, 130 S. Ct. 2705, 2730 (2010) (“We hold that, in regulating the particular forms of support that plaintiffs seek to provide to foreign terrorist organizations [through 18 U.S.C. § 2339B], Congress has pursued that objective consistent with the limitations of the First and Fifth Amendments.”). Thus, any indication that an individual supports an FTO or its political positions could trigger government scrutiny.


87 See Mukasey Guidelines, supra note 37, § II (“These Guidelines do not impose supervisory approval requirements in assessments . . . .”); id. § II.B.2 (providing that “a predicated investigation requires supervisory approval” only if “relating to foreign intelligence”); id. § II.B.5 (requiring FBI Headquarters be notified of predicated investigations only in limited instances); id. (imposing no limits on the duration of investigations).

88 See DIOG, supra note 12, §§ 3.4, 18 (setting supervisory approval requirements, placing time limits on some investigations, and requiring periodic reviews for all investigations); id. § 10 (increasing oversight for “sensitive investigative matters,” such as investigations of politicians, political or religious organizations, or members of the news media).
The 2011 changes to the DIOG loosened some existing restrictions. The DIOG now authorizes a number of investigative techniques even before opening an assessment, such as accessing information in the databases of federal, local, or state governments; interviewing a “complainant”; and searching publicly available information (including social media sites). See Charlie Savage, F.B.I. Agents Get Leeway to Push Privacy Bounds, N.Y. Times, (June 12, 2011), http://www.nytimes.com/2011/06/13/us/13fbi.html (last visited Nov. 13, 2013) (“The new rules add to several measures taken over the past decade to give agents more latitude as they search for signs of criminal or terrorist activity.”) (on file with the Washington and Lee Law Review).

See DIOG, supra note 12, § 2.7.2 (providing the authorization to depart from DIOG procedures). Such departures must not violate the Guidelines themselves. Id.

See Jordan, supra note 86 (“Law enforcement officials say the proposed policy would help them do exactly what Congress demanded after the Sept. 11, 2001, attacks: root out terrorists before they strike.”).

See Frederick A.O. Schwarz, Jr., The Church Committee, Then and Now, in U.S. National Security, Intelligence and Democracy 25 (Russell A. Miller ed., 2008) (“Attorney General Edward Levi and President Gerald Ford, followed by President Jimmy Carter, had issued guidelines and executive orders in response to the Church Committee's revelations that went part of the way toward the goals of the Church Committee.”).


See Schwarz, supra note 92, at 25 (noting the uneasy truce between the FBI's authority and civil liberty).

See Ashcroft Guidelines, supra note 36, § III.B.2 (“The immediate purpose of a terrorism enterprise investigation is to obtain information concerning the nature and structure of the enterprise . . . with a view to the longer range objectives of detection, prevention, and prosecution of the criminal activities of the enterprise.”).

See Am. Civil Liberties Union v. Nat'l Sec. Agency, 493 F.3d 644, 655 (6th Cir. 2007) (noting that the state secrets doctrine would prevent discovery of whether plaintiffs were actually wiretapped).


See infra Part III.C.1 (discussing the deficiencies of administrative self-governance).

See Ashcroft Guidelines, supra note 36, § III (“As a general rule, an investigation of a completed criminal act is normally confined to determining who committed that act and securing evidence to establish the elements of the particular offense.”).

See, e.g., id. § III (recognizing the difference between criminal and intelligence investigations).

Id.

Id. (quoting United States v. U.S. Dist. Ct., 407 U.S. 297, 322 (1972)).

Id. § III.B.3.

See id. § II.B(5)(b), (6)(g) (authorizing the use of nonconsensual electronic surveillance, physical or photographic surveillance, or any other investigative technique covered under chapter 119 of Title 18 of the U.S. Code).

See id. § II.C(1) (“A general crimes investigation may be initiated by the FBI when facts or circumstances reasonably indicate that a federal crime has been, is being, or will be committed.”).

See, e.g., Orin S. Kerr, Applying the Fourth Amendment to the Internet: A General Approach, 62 Stan. L. Rev. 1005, 1013-15 (2010) (recognizing that the digital storage of data means that unlimited data, which can be located anywhere, can be made
available to the government); Privacy and Civil Liberties Oversight Board, Workshop Regarding Surveillance Programs Operated Pursuant to Section 215 of the USA PATRIOT Act and Section 702 of the Foreign Intelligence Surveillance Act, Transcript at 300 (July 9, 2013) (statement of Greg Nojeim, Ctr. for Democracy & Tech.), http://www.pclob.gov/SiteAssets/9-july-2013/Public%20Workshop%20-%20Full.pdf (noting that “something has to substitute for the friction that used to be in the system because there wasn't an ability to collect all this information about all human interaction . . . that we have now”).

See infra notes 121-138 and accompanying text.

See Mukasey Guidelines, supra note 37, § I.D.2 (“These guidelines are not intended to, do not, and may not be relied upon to create any rights, substantive or procedural, enforceable by law by any party in any matter, civil or criminal . . . .”).

See Lujan v. Defenders of Wildlife, 504 U.S. 555, 560 (1992) (“Standing . . . requires the plaintiff to suffer an ‘injury in fact’—an invasion of a legally protected interest which is (a) concrete and particularized, and (b) ‘actual or imminent, not “conjectural” or “hypothetical”’” (footnote omitted)).

See 50 U.S.C. § 1861(d) (2012) (imposing gag orders on entities receiving requests for information from the FBI); 18 U.S.C. § 3103a(b) (allowing law enforcement to delay warrant notice requirements under certain circumstances); Rascoff, supra note 6, at 596 (“[I]ndividuals who allegedly are being spied on illegally tend to be unaware of that fact . . . .” (footnotes omitted)).

See Clapper v. Amnesty Int'l USA, 133 S. Ct. 1138, 1152 (2013) (“Because respondents do not face a threat of certainly impending interception under § 1881a, the costs that they have incurred to avoid surveillance are simply the product of their fear of surveillance . . . [this] is insufficient to create standing.” (footnote omitted)); Laird v. Tatum, 408 U.S. 1, 13-14 (1972) (“Allegations of a subjective ‘chill’ are not an adequate substitute for a claim of specific present objective harm or a threat of specific future harm; the federal courts established pursuant to Article III of the Constitution do not render advisory opinions.”) (quoting United Pub. Workers of Am. (C.I.O.) v. Mitchell, 330 U.S. 75, 89 (1947)); Am. Civil Liberties Union v. Nat'l Sec. Agency, 493 F.3d 644, 660 (6th Cir. 2007) (quoting the “subjective chill” language from Laird); Rascoff, supra note 6, at 596 (“[I]f certain individuals have some basis for thinking that they have been the subjects of illegal surveillance, they are often unable to make . . . [a] definitive showing of injury . . . for constitutional standing.” (footnotes omitted)).

See Mohamed v. Jeppesen Dataplan, Inc., 614 F.3d 1070, 1093 (9th Cir. 2010) (en banc) (dismissing foreign nationals' claims of harm caused by the Central Intelligence Agency's extraordinary rendition program pursuant to the state secrets doctrine); El-Masri v. United States, 479 F.3d 296, 311 (4th Cir. 2007) (dismissing a foreign national's claim of harm caused by the Central Intelligence Agency's extraordinary rendition program pursuant to the state secrets doctrine); Am. Civil Liberties Union v. Nat'l Sec. Agency, 493 F.3d 644, 655 (6th Cir. 2007) (noting that the state secrets doctrine would prevent discovery of whether plaintiffs were actually wiretapped); Fazaga v. FBI, 884 F. Supp. 2d 1022, 1049 (C.D. Cal. 2012) (dismissing claims that the FBI illegally directed agents to gather information on Southern California Muslim Communities pursuant to the state secrets doctrine); Rascoff, supra note 6, at 596 (“[E]ven if [a plaintiff has standing], the government is free to invoke the state secrets privilege and, in effect, unilaterally have the case dismissed . . . .”).

See United States v. Reynolds, 345 U.S. 1, 6-7 (1953) (“[T]he privilege against revealing military secrets . . . is well established in the law of evidence.” (footnote omitted)).

See, e.g., Mohamed v. Jeppesen Dataplan, Inc., 614 F.3d 1070, 1093 (9th Cir. 2010) (en banc) (dismissing foreign nationals' claims of harm caused by the Central Intelligence Agency's extraordinary rendition program pursuant to the state secrets doctrine).

See, e.g., United States v. Felt, 491 F. Supp. 179, 183-84 (D.D.C. 1979) (concluding that in camera, ex-parte review of documents obtained by the FBI from foreign intelligence sources was appropriate).

See Fed. R. Crim. P. 41(h) (providing the ability to file a motion to suppress evidence obtained in violation of the Fourth Amendment).
See United States v. Calandra, 414 U.S. 338, 348 (1974) ("[S] tanding to invoke the exclusionary rule has been confined to situations where the Government seeks to use such evidence to incriminate the victim of the unlawful search.").

See, e.g., Akhil Reed Amar, Fourth Amendment First Principles, 107 Harv. L. Rev. 757, 796-97 (1994) ("Under the exclusionary rule, the more guilty you are, the more you benefit.").

See Calandra, 414 U.S. at 348 (discussing who may challenge government surveillance tactics).

See Schwarz, supra note 92, at 25 (noting early legislative involvement in the Guidelines).


See, e.g., id. at 1-3 (voicing concerns about illegal FBI activities); FBI Statutory Charter Part 2, Hearings Before the Subcomm. on Admin. Practice & Procedure of the S. Comm. on the Judiciary, 95th Cong. 3-4 (1978) (voicing concerns about undercover FBI operations); Schwarz, supra note 92, at 25 ("Attorney General Edward Levi . . . issued guidelines and executive orders in response to the Church Committee's revelations that went part of the way toward the goals of the Church Committee.").

See FBI Statutory Charter Part 1, supra note 122, at 25 (noting that the Guidelines were prompted by the Church Committee's findings and sought to avoid drastic legislative action).

Aziz Z. Huq, Structural Constitutionalism as Counterterrorism, 100 Calif. L. Rev. 887, 921 (2002) ("Errors on the security side are more likely to be widespread, affecting many people and imposing a high political cost.").

See id. at 923 ("[L]egislators will tend . . . to delegate decisions rather than . . . resolve hard questions themselves.").


See U.S. Const. art. I, § 1 (vesting all legislative power in Congress); id. art. I, § 2, cl. 5 (vesting the House of Representatives with the power to impeach public officials); id. art. I, § 3, cl. 6 (vesting the Senate with the power to try all impeachments); id. art. I, § 9, cl. 7 (providing the spending power).

See U.S. Const. art. I, § 9, cl. 7 (providing the spending power).


See McGrain v. Daugherty, 273 U.S. 135, 174 (1927) ("[T]he power of inquiry-with process to enforce it-is an essential and appropriate auxiliary to the legislative function.").

See Rascoff, supra note 6, at 597-98 (providing possible reasons why congressional oversight of intelligence collection lacks “vitality”).

See id. (“Before you sign the guidelines, we urge you to make them available publicly, and to solicit input not only from members of Congress but also from national security and civil liberties experts . . .”).

See Letter from Stephen Kelly, Asst. Dir., FBI Office of Cong. Affairs, to Sen. Richard J. Durbin (July 28, 2010), http://www.bordec.org/press/fbidurbinletter.pdf (alerting the Senator to the fact that the FBI Director “misspoke” when he asserted during an oversight hearing that there is “a requirement of ‘suspicion of wrongdoing’ in order for the FBI to engage in surveillance of an individual or location”).


See 5 U.S.C. § 552(b)(7)(E) (2012) (exempting information that “would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law”).

See supra notes 11-13 and accompanying text. Even if these rules were public, the majority often lacks incentive to object to problematic provisions because much of the burden of intelligence collection falls on minority communities. See David Cole, Enemy Aliens 88-179 (2003) (providing a discussion of the historical pattern of dealing with threats to national security by focusing on noncitizens with the rationale that they have diminished rights).

See supra Part II.A (discussing how the FBI's policies affect the prevention of terrorist attacks and protecting civil liberties).


See Hamed Aleaziz, Want to Sue the FBI for Spying on Your Mosque? Sorry, That's Secret,” Mother Jones (Aug. 8, 2011), http://www.motherjones.com/politics/2011/08/state-secrets-fazaga-v-fbi (last visited Nov. 13, 2013) (quoting University of Texas Law Professor Bobby Chesney's assertion that “[a]t the end of the day, the FBI is part of the intelligence community as well—it's not necessarily thought of as any different than the NSA”) (on file with the Washington and Lee Law Review).

See, e.g., Quick Facts, FBI, supra note 140 (listing civil rights protection as fifth on the FBI’s list of priorities); Intelligence Overview, FBI http://www.fbi.gov/about-us/intelligence/intel-driven/intelligence-overview (last visited Oct. 20, 2013) (discussing “safeguarding civil liberties” last on the “Intelligence Overview” page) (on file with the Washington and Lee Law Review).

Mukasey Guidelines, supra note 37, § I.C.1.

Until 2002, the Guidelines instructed that inquiries and investigations should “be conducted with as little intrusion into the privacy of individuals as the needs of the situation permit.” Memorandum from Charles Doyle, Sr. Specialist, Am. Law Div., Cong. Research Serv., to Senate Select Comm. on Intelligence 13 (Sept. 22, 2008) (citations omitted) (on file with the Washington and Lee Law Review).

Ashcroft Guidelines, supra note 36, § I.

See id. (instructing agents to balance the importance of the information sought with the intrusiveness of the techniques necessary).

See id. § I.C.2. (providing particular guidance relating to terrorism).

See DIOG, supra note 12, Preamble (“To assist the FBI in its mission, the Attorney General signed [DIOG] on September 29, 2008. The primary purpose of the [DIOG] is to standardize policy so that . . . investigative activities are accomplished in a consistent manner.”).
See id. § 4.1.2 (barring the FBI from investigating solely to monitor the exercise of constitutional rights, such as the free exercise of speech, religion, assembly, press and petition” or based “solely on the race, ethnicity, national origin or religious beliefs” of the subject).

Id. § 4.4.4.

Id. § 4.4.5.


See infra Part III.B.

See infra Part III.C.

See infra Part III.D.

See Heather K. Gerken, Shortcuts to Reform, 93 Minn. L. Rev. 1582, 1583-88 (2009) (discussing the difficulty of implementing reforms in the election-administration context and proposing an information-forcing mechanism as a possible catalyst).


See id. at 1825-26 (“[L]egislators can derive scant electoral rewards from associating themselves with particular counterterrorism policy . . . .”).
See id. at 1826 (“Since counterterrorism legislation provides small upside value and enormous downside risk, legislators are best served, electorally, by ensuring that any catastrophe cannot be laid at their feet through voting in . . . a ‘pro-security’ direction.”).


See supra note 31 and accompanying text.

See John Doe, Inc. v. Mukasey, 549 F.3d 861, 876-81 (2d Cir. 2008) (concluding the nondisclosure requirement could not survive strict scrutiny).

See, e.g., Berman, supra note 32, at 34-35 (explaining the importance of community-provided information in FBI counterterrorism efforts).

See, e.g., Ctr. For Human Rights & Global Justice, Targeted & Entrapped: Manufacturing the “Homegrown Threat” in the United States 9-18 (2011) (describing some of the FBI’s post-9/11 information gathering tactics that have attracted scrutiny and criticism).

See, e.g., Office of Mgmt. & Budget, Exec. Office of the President, Final Bulletin for Agency Good Guidance Practices 15 (2007) [hereinafter OMB Bull.] (“As it does for legislative rules, providing pre-adopting opportunity for comment on significant guidance documents can increase the quality of the guidance and provide for greater public confidence in and acceptance of the ultimate agency judgments.”).

See, e.g., Berman, supra note 32, at 43 n.286 (“The Justice Department gave the illusion, though without any substance, of consultation with stakeholders before implementing the current guidelines.”).

See Neal Kumar Katyal, Internal Separation of Powers: Checking Today's Most Dangerous Branch from Within, 115 Yale L.J. 2314, 2323 (2006) (“[S]ometimes broad design choices are easier to impose by fiat than are specific policies.”).


5 U.S.C. § (b)(3)(A) (2012); see also Rascoff, supra note 6, at 644-46 (comparing public participation in the formation of intelligence guidelines with public participation in other agency guidelines).

See Exec. Order No. 12,866, 58 Fed. Reg. 51,735, 51,735 (Sept. 30, 1993) (“Coordinated review of agency rulemaking is necessary . . . The Office of Management and Budget (OMB) shall carry out that review function.”).

Omb Bull., supra note 170, at 1.

See id. (“This Bulletin is intended to increase the quality and transparency of agency guidance practices and the significant guidance documents produced through them.”).
See U.S. Dep't of Homeland Sec., Privacy Impact Assessments: The Privacy Office Official Guidance 4 (2010) [hereinafter PIA Guidance] (“The PIA is a document that helps the public understand what information the Department is collecting . . . . This document builds trust between the public and the Department by increasing transparency of the Department's systems and goals.”).


PIA Guidance, supra note 179, at 1.

See The Federalist Nos. 10, 51 (James Madison) (describing how the Constitution establishes a government conducive to liberty).

U.S. Const. art. 1, § 7, cl. 2.

See Cass R. Sunstein, Is the Clean Air Act Unconstitutional?, 98 Mich. L. Rev. 303, 336 (1999) (“[Congress's] lawmaking power . . . is designed to ensure the combination of deliberation and accountability that comes from saying that government power cannot be brought to bear on individuals unless diverse representatives, from diverse places, have managed to agree on the details.”).

See Jody Freeman, The Private Role in Public Governance, 75 N.Y.U. L. Rev. 543, 545-46 (2000)(“[D]espite their considerable discretionary power to impact individual liberty and property rights, allocate benefits and burdens, and shape virtually every sector of the economy, agencies are not directly accountable to the electorate.

But see Peter H. Schuck, Delegation and Democracy: Comments on David Schoenbrod, 20 Cardozo L. Rev. 775, 783-90 (1999) (outlining ways in which agencies are held democratically accountable).

See e.g., Whitman v. Am. Trucking Ass'ns, Inc., 531 U.S. 457, 487 (2001) (Thomas, J., concurring) (“I believe that there are cases in which the principle is intelligible and yet the significance of the delegated decision is simply too great for the decision to be called anything other than 'legislative.'”); John F. Manning, Textualism and Legislative Intent, 91 Va. L. Rev. 419, 426 n.23 (2005) ("Bicameralism and presentment form an essential component of the constitutional structure, designed to check factional influence, promote caution and deliberation, and provoke public discussion."); Mark Seidenfeld, A Civic Republican Justification for the Bureaucratic State, 105 Harv. L. Rev. 1511, 1512 (1992) ("Over the past century, the powers and responsibilities of administrative agencies have grown to an extent that calls into question the constitutional legitimacy of the modern federal bureaucracy."); Richard B. Stewart, Administrative Law in the Twenty-First Century, 79 N.Y.U. L. Rev. 437, 440 (2003) ("While application of the traditional model might ensure that agencies acted within the bounds of their statutory powers, those bounds were so wide as to give agencies vast discretionary powers, creating a palpable democracy deficit and the threat of arbitrary power.").

See A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495, 542 (1935) (“We think that the code-making authority thus conferred is an unconstitutional delegation of legislative power.”); Panama Refining Co. v. Ryan, 293 U.S. 388, 433 (1935) (invalidating portions of federal legislation for excessive delegation of authority to the executive).

Am. Trucking Ass'n, 531 U.S. at 472 (“[W]e repeatedly have said that when Congress confers decisionmaking authority upon agencies Congress must 'lay down by legislative act an intelligible principle to which the person or body authorized to [act] is directed to conform.'” (citing J.W. Hampton, Jr., & Co. v. United States, 276 U.S. 394, 409 (1928))).

See Kevin M. Stack, The Constitutional Foundations of Chenery, 116 Yale L.J. 952, 958 (2007) (arguing that Chenery's requirement that an agency provide an adequate basis for a rule "operates both to bolster the political accountability of the agency's action and to prevent arbitrariness in the agency's exercise of its discretion"). See generally Cass R. Sunstein, Nondelegation Canons, 67 U. Chi. L. Rev. 315 (2000) (arguing that many judicial rules regarding interpreting the scope of agency power are designed to address the excessive delegation concern).
See Lisa Schultz Bressman, Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State, 78 N.Y.U. L. Rev. 461, 469-92 (2003) (arguing that initial concerns about the administrative state's role in the constitutional structure focused, first and foremost, on the dangers of arbitrary decision making and that such concerns have persisted).

See, e.g., id. at 528 (noting the emergence of reasoned decision making and the “hard look” doctrine); Lisa Schultz Bressman, Procedures as Politics in Administrative Law, 107 Colum. L. Rev. 1749, 1777-79 (2007) (summarizing the history and rationale behind reasoned decision-making requirements in administrative law); Richard B. Stewart, The Reformation of American Administrative Law, 88 Harv. L. Rev. 1667, 1779-80 (1975) (describing judicial insistence that agencies engage in reasoned decision making); see also INS v. Chadha, 462 U.S. 919, 951, 952 n.16 (1983) (noting that the Constitution itself insists that legislation be subjected to deliberation and that the executive's discretion in carrying out legislative mandates is limited by the scope of the legislative delegation).

See Motor Vehicle Mfrs. Ass’n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 41 (1983) (holding that the APA's arbitrary and capricious test requires reasoned decision making); see also, e.g., Smiley v. Citibank (S.D.), N.A., 517 U.S. 735, 741 (1996) (“[W]e have before us here a full-dress regulation, issued by the Comptroller himself and adopted pursuant to the notice-and-comment procedures of the Administrative Procedure Act designed to assure due deliberation . . . .”); Steven P. Croley, Theories of Regulation: Incorporating the Administrative Process, 98 Colum. L. Rev. 1, 79 nn.226-27 (1998) (noting that the APA facilitates deliberative agency decisions); Stewart, supra note 191, at 1670 (noting that APA procedures are “designed to promote the accuracy, rationality, and reviewability of agency application of legislative directives”); Cass R. Sunstein, Interest Groups in American Public Law, 38 Stan. L. Rev. 29, 60-61 (1985) (“Much of modern administrative law is a means of serving the original purposes of the nondelegation doctrine, and of promoting Madisonian goals, without invalidating regulatory statutes or relying on traditional conceptions of private property.”).


See id. at 196 (“[A] reviewing court, in dealing with a determination or judgment which an administrative agency alone is authorized to make, must judge the propriety of such action solely by the grounds invoked by the agency.”); Stack, supra note 189, at 956 (“The [Chenery] principle now applies in review of every form of agency action, from agency rulemaking to informal adjudication, as well as in review of all manner of deficiencies in agency fact-finding and insufficient statements of reasons . . . .”).

The APA lays down procedures for two forms of agency action, rulemaking and adjudication, each of which may be pursued through either formal or informal proceedings. 5 U.S.C. §§ 553, 554, 556, 557.

See Small Refiner Lead Phase-Down Task Force v. EPA, 705 F.2d 506, 547 (D.C. Cir. 1983) (“[N]otice improves the quality of agency rulemaking by ensuring that agency regulations will be ‘tested by exposure to diverse public comment.’” (quoting BASF Wyandotte Corp. v. Costle, 598 F.2d 637, 641 (1st Cir. 1979))).

Id.; Conn. Light & Power Co. v. NRC, 673 F.2d 525, 528 (D.C. Cir. 1982) (“The process of notice and comment rulemaking is . . . to be a process of reasoned decision-making.”).

NRDC v. EPA, 279 F.3d 1180, 1186 (9th Cir. 2002).

Administrative Procedure Act, § 4, 5 U.S.C. § 553 (2012); see also United States v. Nova Scotia Food Prods. Corp., 568 F.2d 240, 251-52 (2d Cir. 1977) (noting that scientific data used in an agency's rulemaking analysis should be disclosed during the notice process); Portland Cement Ass'n v. Ruckelshaus, 486 F.2d 375, 393 (D.C. Cir. 1973) (“It is not consonant with the purpose of a rule-making proceeding to promulgate rules on the basis of . . . data . . . known only to the agency.”).

See id. § 553(c) (“After consideration of the relevant matter presented, the agency shall incorporate in the rules adopted a concise general statement of their basis and purpose.”).
REGULATING DOMESTIC INTELLIGENCE COLLECTION, 71 Wash. & Lee L. Rev. 3

202 Conn. Light & Power Co., 673 F.2d at 528.

203 Independent U.S. Tanker Owners Comm. v. Dole, 809 F.2d 847, 852 (D.C. Cir. 1987); see also S. Rep. No. 752-79, pt. IV(4) (b) (1945) (“The required statement of the basis and purpose of rules issued should not only relate to the data so presented but with reasonable fullness explain the actual basis and objectives of the rule.”).

204 See Stack, supra note 189, at 956 (“The Chenery principle makes the validity of agency action depend upon the validity of contemporaneous agency reason-giving.”).

205 See 5 U.S.C. § 552(b)(7) (2012) (exempting from public disclosure obligations records or information that could reasonably disclose confidential sources, law enforcement procedures or techniques, or lead to circumvention of the law).

206 Neither the Justice Department's Office of Privacy and Civil Liberties nor the Civil Rights Division have been particularly successful in countering the policy preferences of the FBI.

207 See Mukasey Speech, supra note 83 (explaining how the new Attorney General Guidelines would help to transform the FBI into an “elite national security organization” by shifting focus from “investigating crimes after they occur to collecting the intelligence necessary to detect and prevent attacks before they occur”).

208 Ashcroft Guidelines, supra note 36, at 11.

209 See supra Part II.B.

210 See supra Part III.B.

211 See Oversight of the Federal Bureau of Investigation: Hearing Before the S. Comm. on the Judiciary, 110th Cong. 8 (2008) (statement of Robert S. Mueller, Director, FBI) (“[T]he new guidelines are not designed to give the FBI any broad new authorities.”); Mukasey Speech, supra note 83 (“The new consolidated guidelines will, in short, integrate more completely and harmonize the standards that apply to the FBI's activities.”); Press Release, U.S. Dep't of Justice, Briefing with Department Officials on Consolidated Attorney General Guidelines (Sept. 12, 2008) [hereinafter DOJ Briefing] (“[O]n the national security side, it was a new concept to have what were called ‘threat assessments . . . . That was new in 2003; this is not anything different now.”).

212 See DOJ Briefing, supra note 211

And the changed techniques are the physical surveillance, which had been available if you were investigating under general crimes, but not under national security; recruiting and tasking of sources, which again had been available if it was general crimes, but not national security; and then the pretext interviews.

213 Cf. id. (“[T]he decision was reached in 2003 that there needed to be some level of activity before a formal investigation [to] allow the FBI to be proactive. . . . What has changed are some of the techniques that are available in the assessment level.”).

214 See Ashcroft Guidelines, supra note 36, § II.B.6 (listing those investigative techniques available “without any prior authorization from a supervisory agent”).


216 See Bressman, supra note 190, at 542 (explaining that notice-and-comment rulemaking facilitates prospective policymaking including broader perspectives).

217 See supra notes 190-92 and accompanying text.
218 See Bressman, supra note 190, at 542-43 (arguing that notice-and-comment rulemaking leads to a decrease in arbitrariness because of the input from affected parties).

219 See id. at 462 (describing the struggle to reconcile the administrative state with our constitutional structure).

220 See id. (explaining the constant attempt to square the administrative state with a constitutional structure that presumes the accountability of policymakers).

221 See id. at 470-74 (explaining the early models of administrative agencies); Stewart, supra note 186, at 1671-75 (discussing “the Traditional Model” of American administrative law); United States v. Nova Scotia Food Prosds. Corp., 568 F.2d 240, 251-52 (2d Cir. 1977) (discussing the role of scientific data in an agency's rulemaking analysis).

222 See Stewart, supra note 186, at 440-41 (discussing the New Deal regulatory regime which led to the enactment of the Administrative Procedure Act).


225 See, e.g., Croley, supra note 192, at 27 (describing the administrative process as the “proverbial black box that mysteriously translates legislative inputs into regulatory outcomes”); id. at 97 (asserting that the administrative state does not “encourage widespread participation”).

226 See Stewart, supra note 186, at 444 (stating that various methods of “[p]ublic participation . . . have become central foundations of administrative law and practice”; Schuck, supra note 185, at 781 (“Today, the administrative agency is often the site where public participation in lawmaking is most accessible, most meaningful, and most effective.”).

227 See Small Refiner Lead Phase-Down Task Force v. EPA, 705 F.2d 506, 552 (D.C. Cir. 1983) (defining a situation where the EPA's regulations were struck down by the court because the notice of proposed rulemaking was insufficient to put the public fully on notice).

228 See NRDC v. EPA, 279 F.3d 1180, 1186 (9th Cir. 2002) (stating that an agency must give notice to apprise interested parties of the “subjects and issues before the Agency”).

229 See United States v. Nova Scotia Food Prosds., Corp., 568 F.2d 240, 251-52 (2d Cir. 1977) (explaining that an agency that promulgates a rule based on scientific data must make that data available to the public during the comment period).

230 Indep. U.S. Tanker Owners Comm. v. Dole, 809 F.2d 847, 852 (D.C. Cir. 1987); see also S. Doc. No. 79-248, at 20 (2d Sess. 1946) (“The statement of the ‘basis and purpose’ of rules . . . should be fully explanatory of the complete factual and legal basis as well as the object or objects sought.”).

231 See Mark Seidenfeld, Bending the Rules: Flexible Regulation and Constraints on Agency Discretion, 51 Admin. L. Rev. 429, 489 (1999) (discussing how the rulemaking process allows individuals to advance perspectives while others scrutinize those perspectives); see also Conn. Light & Power Co. v. NRC, 673 F.2d 525, 528 (D.C. Cir. 1982) (“One particularly important component of the reasoning process [in rulemaking] is the opportunity for interested parties to participate in a meaningful way. . . . ”).


233 See DIOG, supra note 12, § 4.2.2 (explaining that the FBI may investigate activities or persons so long as it does not infringe on the free exercise of religion guaranteed in the First Amendment).

See DOJ Briefing, supra note 211 (describing the way in which the FBI and DOJ included relevant congressional committees in discussions about consolidating the multiple sets of guidelines into one).

See Rascoff, supra note 6, at 644-48 (explaining that the FBI consulted various groups before the new Attorney General's Guidelines were issued); Savage, supra note 89 (noting that the FBI consulted the ACLU about the new Guidelines prior to their issuance).


See infra notes 329-30.

See J.R. DeShazo & Jody Freeman, Public Agencies as Lobbyists, 105 Colum. L. Rev. 2217, 2221 (2005) (“We argue . . . agencies can be prompted to take their secondary missions more seriously when Congress enhances interagency lobbying by increasing the power of other agencies, which derive relevant expertise and interests from their own statutory mandates, to lobby the implementing agency.” (emphasis added)); Eric Biber, Too Many Things to Do: How to Deal with the Dysfunctions of Multiple-Goal Agencies, 33 Harv. Envt'l L. Rev. 1, 5-6 (2009) (suggesting that in DeShazo & Freeman's model, “the impact of the comments will be based primarily on their persuasiveness or political import, and on the pressure they may place on the decision-making agency to develop better measures of performance on secondary goals”).


DeShazo & Freeman, supra note 239, at 2253.

The Electric Consumers Protection Act (ECPA) required the Federal Energy Regulatory Commission (FERC), rather than the license applicant, to consult with state and federal resource agencies before submitting their applications to FERC, required that FERC establish a dispute resolution process to mediate its disagreements with other agencies, demanded that FERC provide an explanation whenever it chose not to implement the recommendations of other agencies, and forced FERC to engage in monitoring to ensure that dam operators complied with any imposed environmental conditions. See id. at 2225-26 (explaining the added obligations placed on the FERC by the ECPA).

See id. at 2226-27, 2275-80, 2289 (discussing the purpose and effect of the amendment); Biber, supra note 239, at 43 DeShazo and Freeman show through statistical analysis that, after the passage of the statutory changes, FERC consistently imposed more environmental conditions on the approval or renewal of dam licenses, and that this is correlated with increased participation in FERC licensing programs by fish and wildlife agencies.

See 42 U.S.C. § 2000(e) (2012) (establishing the PCLOB and mandating that it “review proposed legislation, regulations, and policies related to efforts to protect the Nation from terrorism”).


See Privacy & Civil Liberties Oversight Bd., Statement (June 20, 2013) (thanking the President for meeting with the board and for providing briefings) (on file with the Washington and Lee Law Review).

See generally Privacy & Civil Liberties Oversight Bd., Workshop Regarding Surveillance Programs Operated Pursuant to Section 215 of the USA PATRIOT Act and Section 702 of the Foreign Intelligence Surveillance Act (2013), http://
www.pclob.gov/All%20Documents/July%209,%C202013%C20Workshop%20Transcript.pdf (seeking suggestions about how to protect civil liberties in light of NSA surveillance programs).

248 See Privacy & Civil Liberties Oversight Bd., Report on the Telephone Records Program Conducted under Section 215 of the USA PATRIOT Act and on the Operations of the Foreign Intelligence Surveillance Court (2013), http://www.pclob.gov/SiteAssets/Pages/default/PCLOB-Report-on-the-Telephone-Records-Program.pdf (analyzing the NSA's telephony metadata collection program and making twelve recommendations for reform); Alston, supra note 245, at 383 (explaining that the PCLOB's task is “to scrutinize privacy and civil liberties issues raised by national security policies and programs”).


250 Any entities involved in decision making regarding the Guidelines' or DIOG's contents also should be entitled to suggest a change to the rules in the same way that the public has the right to petition an agency to issue, modify, or rescind a rule. See 5 U.S.C. § 553(e) (2012) (providing this right).

251 DeShazo & Freeman, supra note 239, at 2220 (“Congress can create the potential for interstatutory conflicts where the agency must balance multiple and potentially competing obligations arising from different statutes usually passed at different times by different enacting majorities.”).

252 See Biber, supra note 239, at 7-8 (pointing out that 16 U.S.C. § 1 and § 668dd(a)(2)-(3)(B) mandate that the Park Service and the Fish and Wildlife Service both conserve and allow for “compatible wildlife-dependent recreation”).

253 Id. at 7.


255 See id. (citing 42 U.S.C. § 4332(B), which requires all federal agencies to consider “unquantified environmental amenities and values” along with “economic and technical considerations,” and § 4332(C), which requires federal agencies to develop environmental impact statements for all major federal actions).

256 See id. at 3 (explaining the problem of multiple-goal agencies and the inevitability that these goals will conflict with each other).

257 See id. at 9 (citing studies and “predict[ing] that agencies faced with conflicting tasks will systematically overperform on the tasks that are easier to measure and have higher incentives, and underperform on the tasks that are harder to measure and have lower incentives.”).

258 See generally Biber, supra note 239 (discussing various strategies agencies use to meet both primary goals and secondary goals). Other commentators have addressed the issue from a more specific angle. See, e.g., DeShazo & Freeman, supra note 239, at 2253 (discussing efforts to ensure that the Federal Energy Regulatory Commission took environmental concerns into account when making licensing decisions); Rachel E. Burkow, Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law, 61 Stan. L. Rev. 869, 869 (2009) (discussing the challenges that arise because prosecutors have the dual role of making both charging and adjudicatory decisions).

259 See Biber, supra note 239, at 32-33 (describing the Wilderness Act, which eliminated several land-management agencies' power to create or eliminate wilderness areas).


262  See supra note 31 and accompanying text.
263  See supra notes 125-35 and accompanying text.
265  See Barkow, supra note 258, at 874 (“The problems posed by federal prosecutors’ combination of adjudicative and enforcement functions are the very same issues raised by the administrative state-and the solutions fit equally well in both settings.”).
267  Id.
269  See Kirshner, supra note 266, at 232-33 (“To underscore its separation from law enforcement, MI5 cannot make arrests or detentions.”).
272  Biber, supra note 239, at 35 (citing, inter alia, William L. Andreen, In Pursuit of NEPA’s Promise: The Role of Executive Oversight in the Implementation of Environmental Policy, 64 Ind. L.J. 205, 205 (1989)).
275  See Brian L. Cole et al., Prospects for Health Impact Assessment in the United States: New and Improved Environmental Impact Assessment or Something Different?, 29 J. Health Pol’y, Pol’y & L. 1153, 1168 (2004) (“NEPA was groundbreaking in that it forced agencies, regardless of their primary mission, to consider the environmental repercussions of their actions . . . .”).
277  Id.
279  See DiMento & Ingram, supra note 274, at 297-98 (arguing that EIAs require “conscious deliberation about the environmental effects of a proposal”).
280  Cole et al., supra note 275, at 1176.
281  See id. (arguing that the EIA process works because it incorporates relevant knowledge at the point of decision making).
See DeShazo & Freeman, supra note 239, at 2222-29 (describing the impact of “interagency lobbying” on FERC’s decision-making process); Biber, supra note 239, at 44 (evaluating the “agency as lobbyist” model).

See supra note 239 and accompanying text.

See supra notes 239-43 and accompanying text (discussing the study and its results).

See Garrett Hatch, Cong. Research Serv., RL34385, Privacy and Civil Liberties Oversight Board: New Independent Agency Status 3 (2012) (suggesting a role for the PCLOB in providing greater oversight for the intelligence community).

See Schlanger, supra note 249, at 47-48 (arguing that for offices dedicated to a particular value to be effective, they must be staffed by individuals who identify themselves professionally as dedicated to that value, rather than to the mission of the agency in which the office is embedded).

See DeShazo & Freeman, supra note 239, at 2261 (noting that agencies might seek to influence each other to achieve a desired outcome through “lobbying,” much like a private lobbyist).

Id. at 2228.


See id. (requiring transmission of a regulatory plan to OIRA for review). But see Jennifer Nou, Agency Self-Insulation Under Presidential Review, 126 Harv. L. Rev. 1755, 1756-64 (2013) (arguing that agencies can effectively insulate their decision making from presidential review).

See DeShazo & Freeman, supra note 239, at 2261 (explaining that third parties may “alert Congress so that it can intervene to correct agency misbehavior”).

See Biber, supra note 239, at 46-50 (explaining how OIRA developed and its role in reviewing the costs of federal agency rules prior to distribution); id. at 48 (“OMB monitors performance of agencies on a secondary goal-maximizing economic efficiency in the achievement of other goals . . .-and requires achievement of at least minimal performance on that goal before it would approve the issuance of a rule.”).

See supra Part III.B.2 (discussing the role of reasoned decision making through notice and comment requirements).

See Schlanger, supra note 249, at 32-33 (discussing how including representatives of a particular point of view in working groups can influence policy).

See id. at 33-35 (detailing how giving an agency clearance authority and the power to conduct reviews are tools that can encourage cooperation with another agency).

Cf. Katyal, supra note 172, at 2316 (arguing that when the “first-best concept” of an executive checked by the legislature is unavailable, “checks and balances must be updated to contemplate second-best executive v. executive divisions”).

See supra note 6 and accompanying text (listing scholars who advocate taking a regulatory approach to the threat of terrorism).

See Stern & Wiener, supra note 6, at 396-97 (arguing that decision makers developing counterterrorism measures “need mechanisms to ensure that sensible risk analysis precedes precautionary actions”).

See Posner, supra note 6, at 690-97 (describing how public fear can hinder government measures and recommending that agencies pay special attention to the “psychology of fear”); Sunstein, supra note 6, at 131-33 (describing some of the problems a government may face in attempting to address the panic that often results from terrorist attacks).

302 See Rascoff, supra note 6, at 617-26 (arguing that centralized review would both increase accuracy and cost-effectiveness of intelligence and protect basic rights); Sunstein, supra note 6, at 131 (advocating cost-benefit analysis of counterterrorism regulations).

303 See Rascoff, supra note 6, at 628-29 (discussing the benefits of judicial review of agency action).

304 See id. at 629-33 (explaining that pluralism, and the related concept of transparency, are beneficial because they provide credibility and additional means through which oversight may occur).

305 Id. at 647.

306 See id. at 616 (“[R]egulatory governance implies a robust framework that simultaneously aims to produce more accurate, [more] cost-effective, and more rights-protecting intelligence.”).

307 See id. at 622-24 (arguing that empowering accountability mechanisms provides the public with a better avenue for protecting rights and restraining abuse).


309 See Rascoff, supra note 6, at 586 (“[T]he FISC ought to provide the sort of judicial review of agency action that I advocate . . . .”).

310 See id. (“The rationality review that I endorse should be performed by an organization within the [ODNI], modeled on [OIRA] within [OMB].”).

311 See supra note 302 and accompanying text (defining and discussing what is meant by the term “rational”).

312 See Rascoff, supra note 6, at 586 (explaining Rascoff's recommended use of the ODNI); Exec. Order No. 12,866, 58 Fed. Reg. 51,735 (Sept. 30, 1993) (stating that an assessment of potential costs and benefits of regulatory action shall be provided to OIRA).

313 See Rascoff, supra note 6, at 625, 634-39 (arguing that rationality review will ensure more effective intelligence collection and provide a secondary benefit to liberty concerns).


315 Rascoff, supra note 6, at 637.


317 See Garrett M. Graff, The Threat Matrix 19 (2011) (describing the ODNI as “staffed with some of the top minds from the FBI, the CIA, Homeland Security, and the Pentagon”); Rascoff, supra note 6, at 637 (noting that the ODNI “possesses the core competences” to discharge its role effectively).


See Rascoff, supra note 6, at 591 (discussing the types of legal rules that apply to intelligence collection and noting the lack of governance for the conduct of human intelligence collection).

See id at 628 (proposing that at some regular interval a court should “review the agency's program for fidelity to the agency’s own stated (and previously approved) objectives”); cf. Bressman, supra note 190, at 529-33 (arguing that agency articulations of the limits on their own power would address concerns about excessive delegations of discretion).


See id. § 1801(b) (defining the term “agent of a foreign power”).

See id. § 1806(e)-(h) (discussing the means by which a person may challenge the use of electronic surveillance in a trial or hearing).

See Rascoff, supra note 6, at 642-44 (explaining that judicial review before the FISC and traditional judicial review differ because FISC review lacks a meaningful adversarial process). Third parties generally lack “incentives . . . to challenge government requests for information.” Id. at 644.

See id. at 639 (discussing the application process for surveillance to the FISC); 50 U.S.C. § 1804 (detailing the requirements of an application for court order to the FISC). In the wake of the NSA surveillance revelations in 2013, some proposals have emerged to inject an adversarial element into FISC proceedings. See, e.g., Matt Sledge, Adam Schiff Prepares FISA Court Bill To Create Special Privacy Advocate, Huffington Post (July 25, 2013, 3:42 PM), http://www.huffingtonpost.com/2013/07/25/ adam-schiff-fisa-court_n_3653946.html (last visited Oct. 19, 2013) (describing one Congressman's plans to introduce legislation creating a special privacy advocate who would appear before the FISC to represent the public interest) (on file with the Washington and Lee Law Review); James G. Carr, A Better Secret Court, N.Y. Times (July 22, 2013), http:// www.nytimes.com/2013/07/23/opinion/a-better-secret-court.html (last visited Oct. 19, 2013) (advocating that FISC judges appoint a lawyer to represent the public interest when novel questions come before it) (on file with the Washington and Lee Law Review). The likely success of such proposals is uncertain; United Kingdom and Canadian efforts in this vein have been roundly criticized. See Joint Comm. on Human Rights, Counter-Terrorism Policy and Human Rights: 28 Days, Intercept and Post-Charge Questioning, 2006-2007, H.L. 157, H.C. 394, at 49-55 (U.K.) (describing the functions of and concerns with the use of Special Advocates to represent the interest of excluded parties in closed hearings).


Rascoff, supra note 6, at 644-66.

See id. at 646 (quoting Senator Russ Feingold asking FBI Director Robert Mueller: “Why can't you at least solicit . . . suggestions in a meaningful process that involves more than a single meeting where the participants aren't even allowed . . . to keep a copy [of the draft guidelines]?”).

Id. at 645.

See supra Part III.D.2 (explaining how the procedural rule changes discussed offer a second-best option).

See Aftergood, supra note 234, at 399 (“[T]he free flow of information to interested members of the public is a prerequisite to their participation in the deliberative process and to their ability to hold elected officials accountable.”).
See Richard J. Pierce, Jr., Seven Ways to Deossify Agency Rulemaking, 47 Admin. L. Rev. 59, 59 (1995) (naming a number of benefits that informal rulemaking confers on society, including the fact that “rulemaking enhances fairness by allowing all potentially affected members of the public to participate in the decision-making process”).


See Fla. Power & Light Co. v. Lorion, 470 U.S. 729, 743-44 (1985) (“The task of the reviewing court is to apply the appropriate APA standard of review to the agency decision based on the record the agency presents to the reviewing court.”).

See Pierce, supra note 334, at 68 (“A principal concern is that without the procedural and substantive requirements of [judicial review], the governing values may be subverted . . . .”); id. at 68 (“[P]rocedural rights are created because of a perception that the existing processes of representation are an inadequate guaranty that the outcome will be something other than the result of private whim.”).

See Citizens to Preserve Overton Park, Inc. v. Volpe, 401 U.S. 402, 419 (1971) (noting that an agency must consider the “whole record” in making a rule, and courts are entitled to review the full administrative record to evaluate the challenged action).

5 U.S.C. § 706(2)(A) (2012); see also Motor Vehicles Mfrs. Ass’n of the U.S. v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 43 (1983) (noting that courts will reject reasoning “so implausible that it could not be ascribed to a difference in view or the product of agency expertise”); Overton Park, 401 U.S. at 416 (explaining that agency decisions reflecting a “clear error of judgment” are unacceptable). Agency decisions are also struck down because the agency “relied on factors” that Congress did not authorize, State Farm, 463 U.S. at 43; the agency failed to consider “relevant factors,” Overton Park, 401 U.S. at 416; or the agency failed to “consider obvious alternatives,” City of Brookings Mun. Tel. Co. v. F.C.C., 822 F.2d 1153, 1169 (D.C. Cir. 1987).

See Pierce, supra note 334, at 68 (noting that judicial review can be beneficial to the extent that “it induces agencies to consider issues and values agencies otherwise would be tempted to ignore”).

See supra note 337-39 and accompanying text (discussing the standard of review imposed on agency actions by the APA).

See supra Part II.B.2.

See supra note 334 and accompanying text (describing the role of the citizen as “watchdog” over agency actions).


See Sinnar, Protecting Rights from Within, supra note 7, at 1047-48 (discussing the Inspector General's review of coercive interrogation techniques used by the CIA). But see id. at 1048-49 (pointing out risks to Inspectors-General independence given their location within an agency and noting that agencies are often reluctant to provide full information to an Inspector General conducting investigations).


See, e.g., Offices of Inspectors Gen. of the Dep't of Def., Dep't of Just., Cent. Intelligence Agency, Nat'l Sec. Agency, Office of the Dir. of Nat'l Intelligence, Unclassified Report on the President's Surveillance Program (2009) (detailing a review of
the President's classified terrorist surveillance program to detect and prevent further attacks on the United States organized after Sept. 11, 2001).


350 See supra Part III (discussing the framework for governance of the administrative state and proposing a governance framework for the rules governing FBI's intelligence collection programs).

351 See supra Part IV (noting that traditional tools for agency governance are not available in the intelligence-collection context and alternative tools may need to be substituted).

352 See supra Part IV.

71 WLLR 3
Public Law 95-511
95th Congress

An Act

To authorize electronic surveillance to obtain foreign intelligence information.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled, That this Act may be cited as the "Foreign Intelligence Surveillance Act of 1978".

TABLE OF CONTENTS

TITLE I—ELECTRONIC SURVEILLANCE WITHIN THE UNITED STATES FOR FOREIGN INTELLIGENCE PURPOSES

| Sec. 101. Definitions. | 50 USC 1801 note. |
| Sec. 102. Authorization for electronic surveillance for foreign intelligence purposes. |
| Sec. 103. Designation of judges. |
| Sec. 104. Application for an order. |
| Sec. 105. Issuance of an order. |
| Sec. 106. Use of information. |
| Sec. 108. Congressional oversight. |
| Sec. 109. Penalties. |
| Sec. 110. Civil liability. |
| Sec. 111. Authorization during time of war. |

TITLE II—CONFORMING AMENDMENTS

Sec. 201. Amendments to chapter 119 of title 18, United States Code.

TITLE III—EFFECTIVE DATE

Sec. 301. Effective date.

DEFINITIONS

Sec. 101. As used in this title:

(a) "Foreign power" means—

(1) a foreign government or any component thereof, whether or not recognized by the United States;
(2) a faction of a foreign nation or nations, not substantially composed of United States persons;
(3) an entity that is openly acknowledged by a foreign government or governments to be directed and controlled by such foreign government or governments;
(4) a group engaged in international terrorism or activities in preparation therefor;
(5) a foreign-based political organization, not substantially composed of United States persons; or
(6) an entity that is directed and controlled by a foreign government or governments.

(b) "Agent of a foreign power" means—

(1) any person other than a United States person, who—
   (A) acts in the United States as an officer or employee of a foreign power, or as a member of a foreign power as defined in subsection (a) (4);
(B) acts for or on behalf of a foreign power which engages in clandestine intelligence activities in the United States contrary to the interests of the United States, when the circumstances of such person's presence in the United States indicate that such person may engage in such activities in the United States, or when such person knowingly aids or abets any person in the conduct of such activities or knowingly conspires with any person to engage in such activities; or
(2) any person who—
(A) knowingly engages in clandestine intelligence gathering activities for or on behalf of a foreign power, which activities involve or may involve a violation of the criminal statutes of the United States;
(B) pursuant to the direction of an intelligence service or network of a foreign power, knowingly engages in any other clandestine intelligence activities for or on behalf of such foreign power, which activities involve or are about to involve a violation of the criminal statutes of the United States;
(C) knowingly engages in sabotage or international terrorism, or activities that are in preparation therefor, for or on behalf of a foreign power; or
(D) knowingly aids or abets any person in the conduct of activities described in subparagraph (A), (B), or (C) or knowingly conspires with any person to engage in activities described in subparagraph (A), (B), or (C).

(c) "International terrorism" means activities that—
(1) involve violent acts or acts dangerous to human life that are a violation of the criminal laws of the United States or of any State, or that would be a criminal violation if committed within the jurisdiction of the United States or any State;
(2) appear to be intended—
(A) to intimidate or coerce a civilian population;
(B) to influence the policy of a government by intimidation or coercion; or
(C) to affect the conduct of a government by assassination or kidnapping; and
(3) occur totally outside the United States, or transcend national boundaries in terms of the means by which they are accomplished, the persons they appear intended to coerce or intimidate, or the locale in which their perpetrators operate or seek asylum.

(d) "Sabotage" means activities that involve a violation of chapter 105 of title 18, United States Code, or that would involve such a violation if committed against the United States.

(e) "Foreign intelligence information" means—
(1) information that relates to, and if concerning a United States person is necessary to, the ability of the United States to protect against—
(A) actual or potential attack or other grave hostile acts of a foreign power or an agent of a foreign power;
(B) sabotage or international terrorism by a foreign power or an agent of a foreign power; or
(C) clandestine intelligence activities by an intelligence service or network of a foreign power or by an agent of a foreign power; or
(2) information with respect to a foreign power or foreign territory that relates to, and if concerning a United States person is necessary to—
   (A) the national defense or the security of the United States; or
   (B) the conduct of the foreign affairs of the United States.

(f) "Electronic surveillance" means—
   (1) the acquisition by an electronic, mechanical, or other surveillance device of the contents of any wire or radio communication sent by or intended to be received by a particular, known United States person who is in the United States, if the contents are acquired by intentionally targeting that United States person, under circumstances in which a person has a reasonable expectation of privacy and a warrant would be required for law enforcement purposes;
   (2) the acquisition by an electronic, mechanical, or other surveillance device of the contents of any wire communication to or from a person in the United States, without the consent of any party thereto, if such acquisition occurs in the United States;
   (3) the intentional acquisition by an electronic, mechanical, or other surveillance device of the contents of any radio communication, under circumstances in which a person has a reasonable expectation of privacy and a warrant would be required for law enforcement purposes, and if both the sender and all intended recipients are located within the United States; or
   (4) the installation or use of an electronic, mechanical, or other surveillance device in the United States for monitoring to acquire information, other than from a wire or radio communication, under circumstances in which a person has a reasonable expectation of privacy and a warrant would be required for law enforcement purposes.

(g) "Attorney General" means the Attorney General of the United States (or Acting Attorney General) or the Deputy Attorney General.

(h) "Minimization procedures", with respect to electronic surveillance, means—
   (1) specific procedures, which shall be adopted by the Attorney General, that are reasonably designed in light of the purpose and technique of the particular surveillance, to minimize the acquisition and retention, and prohibit the dissemination, of nonpublicly available information concerning unconsenting United States persons consistent with the need of the United States to obtain, produce, and disseminate foreign intelligence information;
   (2) procedures that require that nonpublicly available information, which is not foreign intelligence information, as defined in subsection (e)(1), shall not be disseminated in a manner that identifies any United States person, without such person's consent, unless such person's identity is necessary to understand foreign intelligence information or assess its importance;
   (3) notwithstanding paragraphs (1) and (2), procedures that allow for the retention and dissemination of information that is evidence of a crime which has been, is being, or is about...
to be committed and that is to be retained or disseminated for law enforcement purposes; and

(4) notwithstanding paragraphs (1), (2), and (3), with respect to any electronic surveillance approved pursuant to section 102(a), procedures that require that no contents of any communication to which a United States person is a party shall be disclosed, disseminated, or used for any purpose or retained for longer than twenty-four hours unless a court order under section 105 is obtained or unless the Attorney General determines that the information indicates a threat of death or serious bodily harm to any person.

(i) “United States person” means a citizen of the United States, an alien lawfully admitted for permanent residence (as defined in section 101(a)(20) of the Immigration and Nationality Act), an unincorporated association a substantial number of members of which are citizens of the United States or aliens lawfully admitted for permanent residence, or a corporation which is incorporated in the United States, but does not include a corporation or an association which is a foreign power, as defined in subsection (a)(1), (2), or (8).

(j) “United States”, when used in a geographic sense, means all areas under the territorial sovereignty of the United States and the Trust Territory of the Pacific Islands.

(k) “Aggrieved person” means a person who is the target of an electronic surveillance or any other person whose communications or activities were subject to electronic surveillance.

(l) “Wire communication” means any communication while it is being carried by a wire, cable, or other like connection furnished or operated by any person engaged as a common carrier in providing or operating such facilities for the transmission of interstate or foreign communications.

(m) “Person” means any individual, including any officer or employee of the Federal Government, or any group, entity, association, corporation, or foreign power.

(n) “Contents”, when used with respect to a communication, includes any information concerning the identity of the parties to such communication or the existence, substance, purport, or meaning of that communication.

(o) “State” means any State of the United States, the District of Columbia, the Commonwealth of Puerto Rico, the Trust Territory of the Pacific Islands, and any territory or possession of the United States.
(ii) the acquisition of technical intelligence, other than the spoken communications of individuals, from property or premises under the open and exclusive control of a foreign power, as defined in section 101(a) (1), (2), or (3);

(B) there is no substantial likelihood that the surveillance will acquire the contents of any communication to which a United States person is a party; and

(C) the proposed minimization procedures with respect to such surveillance meet the definition of minimization procedures under section 101(h); and

if the Attorney General reports such minimization procedures and any changes thereto to the House Permanent Select Committee on Intelligence and the Senate Select Committee on Intelligence at least thirty days prior to their effective date, unless the Attorney General determines immediate action is required and notifies the committees immediately of such minimization procedures and the reason for their becoming effective immediately.

(2) An electronic surveillance authorized by this subsection may be conducted only in accordance with the Attorney General’s certification and the minimization procedures adopted by him. The Attorney General shall assess compliance with such procedures and shall report such assessments to the House Permanent Select Committee on Intelligence and the Senate Select Committee on Intelligence under the provisions of section 108(a).

(3) The Attorney General shall immediately transmit under seal to the court established under section 103(a) a copy of his certification. Such certification shall be maintained under security measures established by the Chief Justice with the concurrence of the Attorney General, in consultation with the Director of Central Intelligence, and shall remain sealed unless—

(A) an application for a court order with respect to the surveillance is made under sections 101(h)(4) and 104; or

(B) the certification is necessary to determine the legality of the surveillance under section 106(f).

(4) With respect to electronic surveillance authorized by this subsection, the Attorney General may direct a specified communication common carrier to—

(A) furnish all information, facilities, or technical assistance necessary to accomplish the electronic surveillance in such a manner as will protect its secrecy and produce a minimum of interference with the services that such carrier is providing its customers; and

(B) maintain under security procedures approved by the Attorney General and the Director of Central Intelligence any records concerning the surveillance or the aid furnished which such carrier wishes to retain.

The Government shall compensate, at the prevailing rate, such carrier for furnishing such aid.

(b) Applications for a court order under this title are authorized if the President has, by written authorization, empowered the Attorney General to approve applications to the court having jurisdiction under section 103, and a judge to whom an application is made may, notwithstanding any other law, grant an order, in conformity with section 105, approving electronic surveillance of a foreign power or an agent of a foreign power for the purpose of obtaining foreign intelligence information, except that the court shall not have jurisdic-
Court to hear applications and grant orders. 50 USC 1803.

Sec. 103. (a) The Chief Justice of the United States shall publicly designate seven district court judges from seven of the United States judicial circuits who shall constitute a court which shall have jurisdiction to hear applications for and grant orders approving electronic surveillance anywhere within the United States under the procedures set forth in this Act, except that no judge designated under this subsection shall hear the same application for electronic surveillance under this Act which has been denied previously by another judge designated under this subsection. If any judge so designated denies an application for an order authorizing electronic surveillance under this Act, such judge shall provide immediately for the record a written statement of each reason for his decision and, on motion of the United States, the record shall be transmitted, under seal, to the court of review established in subsection (b).

(b) The Chief Justice shall publicly designate three judges, one of whom shall be publicly designated as the presiding judge, from the United States district courts or courts of appeals who together shall comprise a court of review which shall have jurisdiction to review the denial of any application made under this Act. If such court determines that the application was properly denied, the court shall immediately provide for the record a written statement of each reason for its decision and, on petition of the United States for a writ of certiorari, the record shall be transmitted under seal to the Supreme Court, which shall have jurisdiction to review such decision.

(c) Proceedings under this Act shall be conducted as expeditiously as possible. The record of proceedings under this Act, including applications made and orders granted, shall be maintained under security measures established by the Chief Justice in consultation with the Attorney General and the Director of Central Intelligence.

(d) Each judge designated under this section shall so serve for a maximum of seven years and shall not be eligible for redesignation, except that the judges first designated under subsection (a) shall be designated for terms of from one to seven years so that one term expires each year, and that judges first designated under subsection (b) shall be designated for terms of three, five, and seven years.

APPLICATION FOR AN ORDER

Approval of Attorney General. 50 USC 1804.

Sec. 104. (a) Each application for an order approving electronic surveillance under this title shall be made by a Federal officer in writing upon oath or affirmation to a judge having jurisdiction under section 103. Each application shall require the approval of the Attorney General based upon his finding that it satisfies the criteria and requirements of such application as set forth in this title. It shall include—

(1) the identity of the Federal officer making the application;
(2) the authority conferred on the Attorney General by the President of the United States and the approval of the Attorney General to make the application;
(3) the identity, if known, or a description of the target of the electronic surveillance;
(4) a statement of the facts and circumstances relied upon by the applicant to justify his belief that—
   (A) the target of the electronic surveillance is a foreign power or an agent of a foreign power; and
   (B) each of the facilities or places at which the electronic surveillance is directed is being used, or is about to be used, by a foreign power or an agent of a foreign power;
(5) a statement of the proposed minimization procedures;
(6) a detailed description of the nature of the information sought and the type of communications or activities to be subjected to the surveillance;
(7) a certification or certifications by the Assistant to the President, for National Security Affairs or an executive branch official or officials designated by the President from among those executive officers employed in the area of national security or defense and appointed by the President with the advice and consent of the Senate—
   (A) that the certifying official deems the information sought to be foreign intelligence information;
   (B) that the purpose of the surveillance is to obtain foreign intelligence information;
   (C) that such information cannot reasonably be obtained by normal investigative techniques;
   (D) that designates the type of foreign intelligence information being sought according to the categories described in section 101(e); and
   (E) including a statement of the basis for the certification that—
      (i) the information sought is the type of foreign intelligence information designated; and
      (ii) such information cannot reasonably be obtained by normal investigative techniques;
(8) a statement of the means by which the surveillance will be effected and a statement whether physical entry is required to effect the surveillance;
(9) a statement of the facts concerning all previous applications that have been made to any judge under this title involving any of the persons, facilities, or places specified in the application, and the action taken on each previous application;
(10) a statement of the period of time for which the electronic surveillance is required to be maintained, and if the nature of the intelligence gathering is such that the approval of the use of electronic surveillance under this title should not automatically terminate when the described type of information has first been obtained, a description of facts supporting the belief that additional information of the same type will be obtained thereafter; and
(11) whenever more than one electronic, mechanical or other surveillance device is to be used with respect to a particular proposed electronic surveillance, the coverage of the devices involved and what minimization procedures apply to information acquired by each device.

(b) Whenever the target of the electronic surveillance is a foreign power, as defined in section 101(a) (1), (2), or (3), and each of the facilities or places at which the surveillance is directed is owned, leased, or exclusively used by that foreign power, the application need not contain the information required by paragraphs (6), (7) (E), (8),
and (11) of subsection (a), but shall state whether physical entry is
required to effect the surveillance and shall contain such information
about the surveillance techniques and communications or other
information concerning United States persons likely to be obtained as
may be necessary to assess the proposed minimization procedures.

(c) The Attorney General may require any other affidavit or
certification from any other officer in connection with the application.

(d) The judge may require the applicant to furnish such other
information as may be necessary to make the determinations required
by section 105.

ISSUANCE OF AN ORDER

SEC. 105. (a) Upon an application made pursuant to section 104,
the judge shall enter an ex parte order as requested or as modified
approving the electronic surveillance if he finds that—

(1) the President has authorized the Attorney General to
approve applications for electronic surveillance for foreign
intelligence information;

(2) the application has been made by a Federal officer and
approved by the Attorney General;

(3) on the basis of the facts submitted by the applicant there is
probable cause to believe that—

(A) the target of the electronic surveillance is a foreign
power or an agent of a foreign power: Provided, That no
United States person may be considered a foreign power or
an agent of a foreign power solely upon the basis of activities
protected by the first amendment to the Constitution of the
United States; and

(B) each of the facilities or places at which the electronic
surveillance is directed is being used, or is about to be used,
by a foreign power or an agent of a foreign power;

(4) the proposed minimization procedures meet the definition
of minimization procedures under section 101(h); and

(5) the application which has been filed contains all statements
and certifications required by section 104 and, if the target is a
United States person, the certification or certifications are not
clearly erroneous on the basis of the statement made under section
104(a)(7)(E) and any other information furnished under section
104(d).

(b) An order approving an electronic surveillance under this section
shall—

(1) specify—

(A) the identity, if known, or a description of the target of
the electronic surveillance;

(B) the nature and location of each of the facilities or
places at which the electronic surveillance will be directed;

(C) the type of information sought to be acquired and the
type of communications or activities to be subjected to the
surveillance;

(D) the means by which the electronic surveillance will be
effected and whether physical entry will be used to effect the
surveillance;

(E) the period of time during which the electronic surveil-
ance is approved; and

(F) whenever more than one electronic, mechanical, or
other surveillance device is to be used under the order, the
authorized coverage of the devices involved and what minimi-
(2) direct—
   (A) that the minimization procedures be followed;
   (B) that, upon the request of the applicant, a specified communication or other common carrier, landlord, custodian, or other specified person furnish the applicant forthwith all information, facilities, or technical assistance necessary to accomplish the electronic surveillance in such a manner as will protect its secrecy and produce a minimum of interference with the services that such carrier, landlord, custodian, or other person is providing that target of electronic surveillance;
   (C) that such carrier, landlord, custodian, or other person maintain under security procedures approved by the Attorney General and the Director of Central Intelligence any records concerning the surveillance or the aid furnished that such person wishes to retain; and
   (D) that the applicant compensate, at the prevailing rate, such carrier, landlord, custodian, or other person for furnishing such aid.

(c) Whenever the target of the electronic surveillance is a foreign power, as defined in section 101(a) (1), (2), or (3), and each of the facilities or places at which the surveillance is directed is owned, leased, or exclusively used by that foreign power, the order need not contain the information required by subparagraphs (C), (D), and (F) of subsection (b)(1), but shall generally describe the information sought, the communications or activities to be subjected to the surveillance, and the type of electronic surveillance involved, including whether physical entry is required.

(d) (1) An order issued under this section may approve an electronic surveillance for the period necessary to achieve its purpose, or for ninety days, whichever is less, except that an order under this section shall approve an electronic surveillance targeted against a foreign power, as defined in section 101(a) (1), (2), or (3), for the period specified in the application or for one year, whichever is less.
   (2) Extensions of an order issued under this title may be granted on the same basis as an original order upon an application for an extension and new findings made in the same manner as required for an original order, except that an extension of an order under this Act for a surveillance targeted against a foreign power, as defined in section 101(a) (3) or (6), or against a foreign power as defined in section 101(a) (4) that is not a United States person, may be for a period not to exceed one year if the judge finds probable cause to believe that no communication of any individual United States person will be acquired during the period.
   (3) At or before the end of the period of time for which electronic surveillance is approved by an order or an extension, the judge may assess compliance with the minimization procedures by reviewing the circumstances under which information concerning United States persons was acquired, retained, or disseminated.
   (e) Notwithstanding any other provision of this title, when the Attorney General reasonably determines that—
      (1) an emergency situation exists with respect to the employment of electronic surveillance to obtain foreign intelligence information before an order authorizing such surveillance can with due diligence be obtained; and
Emergency order. He may authorize the emergency employment of electronic surveillance if a judge having jurisdiction under section 103 is informed by the Attorney General or his designee at the time of such authorization that the decision has been made to employ emergency electronic surveillance and if an application in accordance with this title is made to that judge as soon as practicable, but not more than twenty-four hours after the Attorney General authorizes such surveillance. If the Attorney General authorizes such emergency employment of electronic surveillance, he shall require that the minimization procedures required by this title for the issuance of a judicial order be followed. In the absence of a judicial order approving such electronic surveillance, the surveillance shall terminate when the information sought is obtained, when the application for the order is denied, or after the expiration of twenty-four hours from the time of authorization by the Attorney General, whichever is earliest. In the event that such application for approval is denied, or in any other case where the electronic surveillance is terminated and no order is issued approving the surveillance, no information obtained or evidence derived from such surveillance shall be received in evidence or otherwise disclosed in any trial, hearing, or other proceeding in or before any court, grand jury, department, office, agency, regulatory body, legislative committee, or other authority of the United States, a State, or political subdivision thereof, and no information concerning any United States person acquired from such surveillance shall subsequently be used or disclosed in any other manner by Federal officers or employees without the consent of such person, except with the approval of the Attorney General if the information indicates a threat of death or serious bodily harm to any person. A denial of the application made under this subsection may be reviewed as provided in section 103.

(f) Notwithstanding any other provision of this title, officers, employees, or agents of the United States are authorized in the normal course of their official duties to conduct electronic surveillance not targeted against the communications of any particular person or persons, under procedures approved by the Attorney General, solely to—

(1) test the capability of electronic equipment, if—

   (A) it is not reasonable to obtain the consent of the persons incidentally subjected to the surveillance;
   (B) the test is limited in extent and duration to that necessary to determine the capability of the equipment;
   (C) the contents of any communication acquired are retained and used only for the purpose of determining the capability of the equipment, are disclosed only to test personnel, and are destroyed before or immediately upon completion of the test; and:
   (D) Provided, That the test may exceed ninety days only with the prior approval of the Attorney General;

(2) determine the existence and capability of electronic surveillance equipment being used by persons not authorized to conduct electronic surveillance, if—

   (A) it is not reasonable to obtain the consent of persons incidentally subjected to the surveillance;
   (B) such electronic surveillance is limited in extent and duration to that necessary to determine the existence and capability of such equipment; and
(C) any information acquired by such surveillance is used only to enforce chapter 119 of title 18, United States Code, or section 605 of the Communications Act of 1934, or to protect information from unauthorized surveillance; or
(3) train intelligence personnel in the use of electronic surveillance equipment, if—
(A) it is not reasonable to—
(i) obtain the consent of the persons incidentally subjected to the surveillance;
(ii) train persons in the course of surveillances otherwise authorized by this title; or
(iii) train persons in the use of such equipment without engaging in electronic surveillance;
(B) such electronic surveillance is limited in extent and duration to that necessary to train the personnel in the use of the equipment; and
(C) no contents of any communication acquired are retained or disseminated for any purpose, but are destroyed as soon as reasonably possible.

(g) Certifications made by the Attorney General pursuant to section 102(a) and applications made and orders granted under this title shall be retained for a period of at least ten years from the date of the certification or application.

USE OF INFORMATION

Sec. 106. (a) Information acquired from an electronic surveillance conducted pursuant to this title concerning any United States person may be used and disclosed by Federal officers and employees without the consent of the United States person only in accordance with the minimization procedures required by this title. No otherwise privileged communication obtained in accordance with, or in violation of, the provisions of this title shall lose its privileged character. No information acquired from an electronic surveillance pursuant to this title may be used or disclosed by Federal officers or employees except for lawful purposes.
(b) No information acquired pursuant to this title shall be disclosed for law enforcement purposes unless such disclosure is accompanied by a statement that such information, or any information derived therefrom, may only be used in a criminal proceeding with the advance authorization of the Attorney General.
(c) Whenever the Government intends to enter into evidence or otherwise use or disclose in any trial, hearing, or other proceeding in or before any court, department, officer, agency, regulatory body, or other authority of the United States, against an aggrieved person, any information obtained or derived from an electronic surveillance of that aggrieved person pursuant to the authority of this title, the Government shall, prior to the trial, hearing, or other proceeding or at a reasonable time prior to an effort to so disclose or so use that information or submit it in evidence, notify the aggrieved person and the court or other authority in which the information is to be disclosed or used that the Government intends to so disclose or so use such information.
(d) Whenever any State or political subdivision thereof intends to enter into evidence or otherwise use or disclose in any trial, hearing, or other proceeding in or before any court, department, officer, agency, regulatory body, or other authority of a State or a political subdivision thereof, against an aggrieved person any information obtained or
derived from an electronic surveillance of that aggrieved person pursuant to the authority of this title, the State or political subdivision thereof shall notify the aggrieved person, the court or other authority in which the information is to be disclosed or used, and the Attorney General that the State or political subdivision thereof intends to so disclose or so use such information.

(e) Any person against whom evidence obtained or derived from an electronic surveillance to which he is an aggrieved person is to be, or has been, introduced or otherwise used or disclosed in any trial, hearing, or other proceeding in or before any court, department, officer, agency, regulatory body, or other authority of the United States, a State, or a political subdivision thereof, may move to suppress the evidence obtained or derived from such electronic surveillance on the grounds that—

(1) the information was unlawfully acquired; or
(2) the surveillance was not made in conformity with an order of authorization or approval.

Such a motion shall be made before the trial, hearing, or other proceeding unless there was no opportunity to make such a motion or the person was not aware of the grounds of the motion.

(f) Whenever a court or other authority is notified pursuant to subsection (c) or (d), or whenever a motion is made pursuant to subsection (e), or whenever any motion or request is made by an aggrieved person pursuant to any other statute or rule of the United States or any State before any court or other authority of the United States or any State to discover or obtain applications or orders or other materials relating to electronic surveillance or to discover, obtain, or suppress evidence or information obtained or derived from electronic surveillance under this Act, the United States district court or, where the motion is made before another authority, the United States district court in the same district as the authority, shall, notwithstanding any other law, if the Attorney General files an affidavit under oath that disclosure or an adversary hearing would harm the national security of the United States, review in camera and ex parte the application, order, and such other materials relating to the surveillance as may be necessary to determine whether the surveillance of the aggrieved person was lawfully authorized and conducted. In making this determination, the court may disclose to the aggrieved person, under appropriate security procedures and protective orders, portions of the application, order, and such other materials relating to the surveillance only where such disclosure is necessary to make an accurate determination of the legality of the surveillance.

(g) If the United States district court pursuant to subsection (f) determines that the surveillance was not lawfully authorized or conducted, it shall, in accordance with the requirements of law, suppress the evidence which was unlawfully obtained or derived from electronic surveillance of the aggrieved person or otherwise grant the motion of the aggrieved person. If the court determines that the surveillance was lawfully authorized and conducted, it shall deny the motion of the aggrieved person except to the extent that due process requires discovery or disclosure.

(h) Orders granting motions or requests under subsection (g), decisions under this section that electronic surveillance was not lawfully authorized or conducted, and orders of the United States district court requiring review or granting disclosure of applications, orders, or other materials relating to a surveillance shall be final orders and binding upon all courts of the United States and the several States
except a United States court of appeals and the Supreme Court.

(i) In circumstances involving the unintentional acquisition by an electronic, mechanical, or other surveillance device of the contents of any radio communication, under circumstances in which a person has a reasonable expectation of privacy and a warrant would be required for law enforcement purposes, and if both the sender and all intended recipients are located within the United States, such contents shall be destroyed upon recognition, unless the Attorney General determines that the contents indicate a threat of death or serious bodily harm to any person.

(j) If an emergency employment of electronic surveillance is authorized under section 105(e) and a subsequent order approving the surveillance is not obtained, the judge shall cause to be served on any United States person named in the application and on such other United States persons subject to electronic surveillance as the judge may determine in his discretion it is in the interest of justice to serve, notice of—

(1) the fact of the application;
(2) the period of the surveillance; and
(3) the fact that during the period information was or was not obtained.

On an ex parte showing of good cause to the judge the serving of the notice required by this subsection may be postponed or suspended for a period not to exceed ninety days. Thereafter, on a further ex parte showing of good cause, the court shall forego ordering the serving of the notice required under this subsection.

Disposal of contents.

Postponement or suspension of notice, time limitation.

REPORT OF ELECTRONIC SURVEILLANCE

Sec. 107. In April of each year, the Attorney General shall transmit to the Administrative Office of the United States Court and to Congress a report setting forth with respect to the preceding calendar year—

(a) the total number of applications made for orders and extensions of orders approving electronic surveillance under this title; and
(b) the total number of such orders and extensions either granted, modified, or denied.

Report to Congress. 50 USC 1807.

CONGRESSIONAL OVERSIGHT

Sec. 108. (a) On a semianual basis the Attorney General shall fully inform the House Permanent Select Committee on Intelligence and the Senate Select Committee on Intelligence concerning all electronic surveillance under this title. Nothing in this title shall be deemed to limit the authority and responsibility of the appropriate committees of each House of Congress to obtain such information as they may need to carry out their respective functions and duties.

(b) On or before one year after the effective date of this Act and on the same day each year for four years thereafter, the Permanent Select Committee on Intelligence and the Senate Select Committee on Intelligence shall report respectively to the House of Representatives and the Senate, concerning the implementation of this Act. Said reports shall include but not be limited to an analysis and recommendations concerning whether this Act should be (1) amended, (2) repealed, or (3) permitted to continue in effect without amendment.

Report to congressional committees. 50 USC 1808.

Report of congressional committees to Congress.
Penalties

50 U.S.C. 1809. SEC. 109. (a) Offense.—A person is guilty of an offense if he intentionally—

1. engages in electronic surveillance under color of law except as authorized by statute; or
2. discloses or uses information obtained under color of law by electronic surveillance, knowing or having reason to know that the information was obtained through electronic surveillance not authorized by statute.

(b) Defense.—It is a defense to a prosecution under subsection (a) that the defendant was a law enforcement or investigative officer engaged in the course of his official duties and the electronic surveillance was authorized by and conducted pursuant to a search warrant or court order of a court of competent jurisdiction.

(c) Penalty.—An offense described in this section is punishable by a fine of not more than $10,000 or imprisonment for not more than five years, or both.

(d) Jurisdiction.—There is Federal jurisdiction over an offense under this section if the person committing the offense was an officer or employee of the United States at the time the offense was committed.

Civil Liability

50 U.S.C. 1810. SEC. 110. Civil Action.—An aggrieved person, other than a foreign power or an agent of a foreign power, as defined in section 101 (a) or (b) (1) (A), respectively, who has been subjected to an electronic surveillance or about whom information obtained by electronic surveillance of such person has been disclosed or used in violation of section 109 shall have a cause of action against any person who committed such violation and shall be entitled to recover—

(a) actual damages, but not less than liquidated damages of $1,000 or $100 per day for each day of violation, whichever is greater;
(b) punitive damages; and
(c) reasonable attorney’s fees and other investigation and litigation costs reasonably incurred.

Authorization During Time of War

50 U.S.C. 1811. SEC. 111. Notwithstanding any other law, the President, through the Attorney General, may authorize electronic surveillance without a court order under this title to acquire foreign intelligence information for a period not to exceed fifteen calendar days following a declaration of war by the Congress.

Title II—Conforming Amendments

Amendments to Chapter 119 of Title 18, United States Code

Sec. 201. Chapter 119 of title 18, United States Code, is amended as follows:

18 U.S.C. 2511. (a) Section 2511 (2) (a) (ii) is amended to read as follows:
"(ii) Notwithstanding any other law, communication common carriers, their officers, employees, and agents, landlords, custodians, or other persons, are authorized to provide information, facilities, or technical assistance to persons authorized by law to intercept wire or
oral communications or to conduct electronic surveillance, as defined in section 101 of the Foreign Intelligence Surveillance Act of 1978, if the common carrier, its officers, employees, or agents, landlord, custodian, or other specified person, has been provided with—

"(A) a court order directing such assistance signed by the authorizing judge, or

"(B) a certification in writing by a person specified in section 2518(7) of this title or the Attorney General of the United States that no warrant or court order is required by law, that all statutory requirements have been met, and that the specified assistance is required,

setting forth the period of time during which the provision of the information, facilities, or technical assistance is authorized and specifying the information, facilities, or technical assistance required. No communication common carrier, officer, employee, or agent thereof, or landlord, custodian, or other specified person shall disclose the existence of any interception or surveillance or the device used to accomplish the interception or surveillance with respect to which the person has been furnished an order or certification under this subparagraph, except as may otherwise be required by legal process and then only after prior notification to the Attorney General or to the principal prosecuting attorney of a State or any political subdivision of a State, as may be appropriate. Any violation of this subparagraph by a communication common carrier or an officer, employee, or agent thereof, shall render the carrier liable for the civil damages provided for in section 2520. No cause of action shall lie in any court against any communication common carrier, its officers, employees, or agents, landlord, custodian, or other specified person for providing information, facilities, or assistance in accordance with the terms of an order or certification under this subparagraph."

(b) Section 2511(2) is amended by adding at the end thereof the following new provisions:

"(e) Notwithstanding any other provision of this title or section 605 or 606 of the Communications Act of 1934, it shall not be unlawful for an officer, employee, or agent of the United States in the normal course of his official duty to conduct electronic surveillance, as defined in section 101 of the Foreign Intelligence Surveillance Act of 1978, as authorized by that Act.

"(f) Nothing contained in this chapter, or section 605 of the Communications Act of 1934, shall be deemed to affect the acquisition by the United States Government of foreign intelligence information from international or foreign communications by a means other than electronic surveillance as defined in section 101 of the Foreign Intelligence Surveillance Act of 1978, and procedures in this chapter and the Foreign Intelligence Surveillance Act of 1978 shall be the exclusive means by which electronic surveillance, as defined in section 101 of such Act, and the interception of domestic wire and oral communications may be conducted."

(c) Section 2511(3) is repealed.

(d) Section 2518(1) is amended by inserting "under this chapter" after "communication".

(e) Section 2518(4) is amended by inserting "under this chapter" after both appearances of "wire or oral communication".

(f) Section 2518(9) is amended by striking out "intercepted" and inserting "intercepted pursuant to this chapter" after "communication".
(g) Section 2518(10) is amended by striking out "intercepted" and inserting "intercepted pursuant to this chapter" after the first appearance of "communication".

(h) Section 2519(3) is amended by inserting "pursuant to this chapter" after "wire or oral communications" and after "granted or denied".

**TITLE III—EFFECTIVE DATE**

**EFFECTIVE DATE**

SEC. 301. The provisions of this Act and the amendments made hereby shall become effective upon the date of enactment of this Act, except that any electronic surveillance approved by the Attorney General to gather foreign intelligence information shall not be deemed unlawful for failure to follow the procedures of this Act, if that surveillance is terminated or an order approving that surveillance is obtained under title I of this Act within ninety days following the designation of the first judge pursuant to section 103 of this Act.

Panel IV:

Cyber Operations and the Law of Armed Conflict

Moderator:
Col. David E. Graham

Discussants
Col. Gary Brown
Col. Gary Corn
Rebecca Crootof
Maj. Gen. Charles Dunlap
INTERNATIONAL LAW APPLIES TO CYBER WARFARE! NOW WHAT?

Gary D. Brown*

INTRODUCTION

It’s no longer controversial (if it ever was) to say international law applies to cyber warfare. The United Nations (UN) has said “[i]nternational law, and in particular the Charter of the United Nations, is applicable.”¹ State Department Legal Adviser Harold Koh expressed existing U.S. policy in 2012 when he officially stated that “international law principles do apply in cyberspace.”² And, expressing the unanimous view of the international group of experts gathered to develop the first comprehensive text on cyber international law, Rule 80 of the Tallinn Manual on the International Law Applicable to Cyber Warfare (Tallinn Manual)—which gives away the ending with the title—notes that international law applies to cyber warfare.³

So, yes, international law applies to cyber warfare. But international law relevant to warfare comes in two flavors, as Harold Koh noted:

Under international law, there are two distinct ways of looking at war—the reasons you fight and how you fight. In theory, it is possible to break all the rules while fighting a just war or to be engaged in an unjust war while adhering to the laws of armed conflict. For this reason, the two branches of law are completely independent of one another.

* Professor of Cyber Security, Marine Corps University, Quantico, Virginia.

³. TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE 375 (2d ed. 2017).
Jus (or ius) ad bellum is the title given to the branch of law that defines the legitimate reasons a state may engage in war and focuses on certain criteria that render a war just. . . . Jus in bello, by contrast, is the set of laws that come into effect once a war has begun.  

In contrast to the critical role played by *jus ad bellum*, this article suggests the *jus in bello* (law of armed conflict or LOAC) has little of interest to say specifically about cyber warfare. Mr. Koh goes on to note specifically that the law of armed conflict applies in the context of armed conflict, and the principles of distinction, proportionality, and necessity are applicable in that context. Although the legal overlap is significant, the places where actual cyber activities and the LOAC intersect are few and, in the scheme of things, not especially relevant. What is more, it could be damaging to attempt to flex LOAC to cover cyber operations not within the meaning of the law.

Again, this is not to say law is irrelevant to cyber operations. Law is a critical aspect of discussions about the cyber aspects of privacy rights, espionage, sovereignty, international norms of behavior, and more—but none of these things is within the realm of LOAC. LOAC applies only during armed conflict. In addition to there being ambiguity about what pure cyber armed conflict would look like, there are issues with how relevant LOAC would be regarding the use of cyber techniques in the context of traditional armed conflict when the techniques do not result in kinetic effects.

Set out below are observations of what makes cyber conflict unique, followed by a discussion of law other than LOAC that is relevant to cyber operations and a case for an effects-based evaluation of cyber operations. The paper concludes with a look at why, for practical reasons, LOAC as a body of law has little relevance for cyber warfare, and the danger in trying to interpret it creatively enough to make it matter for cyber operations in armed conflicts.

---


5. The law applicable during armed conflict is referred to as both International Humanitarian Law (IHL) and the Law of Armed Conflict (LOAC). They refer to same body of law and are used interchangeably throughout this paper.


7. Of course, cyber activities that result in death or destruction would be analyzed according to the effect, just as bombs and bullets are. These events are not the subject of this article, and are analyzed precisely the same as their kinetic counterparts.
2017] INTERNATIONAL LAW APPLIES TO CYBER WARFARE! 357

UNIQUENESS OF CYBER WARFARE

Despite assertions to the contrary, cyber-based warfare is *a lot different* from traditional kinetic warfare. In the past, the introduction of new technologies into warfare hasn’t caused LOAC to break a sweat. It has been straightforward to apply traditional law to situations in which violence in warfare has been carried out by a new method. However armed conflict has been conducted, there haven’t been significant debates about whether a given capability somehow eluded being governed by LOAC, although there have been issues around the edges about how LOAC would be applied.

For example, airpower was introduced as a means of warfare in the 20th century but, even though it was new in many ways, it did nothing to challenge experts’ intuitive understanding of warfare. Airpower still employed kinetic munitions, just like artillery and naval guns, both of which had been around for many years. The same basic rules applied. Later, precision-guided munitions (PGMs) were introduced but, again, there was not really anything new there. PGMs are simply more accurate than dumb bombs, creating some debate over whether their use is mandatory when they are available, but not controversy over whether LOAC governs explosions caused by PGMs.

The argument about nuclear weapons has generally focused on whether LOAC bans them entirely as indiscriminate, not whether the body of law controls their use in armed conflict.

These new methods did little to interrupt the functioning of the *jus ad bellum*, either. Explosives and ballistic munitions, however delivered, are similar in effect. Whether a crater is caused by artillery or an air-delivered munition is of but little relevance when considering whether it constitutes an

---


10. *See generally Rid, supra note 9.


armed attack. The analogies between territory and airspace, and space and the high seas, are strong and permit fairly straightforward solutions to the most vexing issues regarding sovereignty.  

Cyber armed conflict, on the other hand, has introduced a host of unique issues to the bodies of international law governing warfare. Time and geography offer few limits to cyber operations, which can happen in less than the blink of an eye anywhere on the globe. Further, most of the modern LOAC developed when States had a monopoly on the means of warfare but, unlike tanks, ships, and bombers, cyber techniques are widely available to the public. Also, there are real questions about which cyberspace activities would violate “cyber sovereignty.” For example, do electronic penetrations of computer systems violate territorial sovereignty as military invasions do?

Another difficult issue is that the infrastructure that is used to carry out legitimate and important civilian business and education is the same infrastructure used to engage in cyber espionage, carry out cyber aggression, to conduct strategic communications, and to do just about everything of importance a State government or its population would do. This may have an unfortunate practical effect on the notion of protecting civilian infrastructure, because there really is no purely civilian cyber infrastructure. The commingling of military/security and civilian infrastructure tends to make the principle of distinction less relevant, if not altogether academic.

Perhaps the biggest issue facing States as they puzzle through how to govern cyber warfare is that crime, espionage, and warfare in cyberspace are all identical to a point. Unlike kinetic operations, which are different in kind and scale from crime and espionage, cyber warfare operations can be utterly indistinguishable from cyber crime and peacetime cyber espionage. This creates new issues for States trying to determine how they may, and how they should, react to adversary cyber operations they discover ongoing.

Despite the unique qualities of cyber capabilities, there should be no confusion about whether LOAC applies to cyber warfare—it does. There is no exception that would exempt cyber warfare from being governed by


LOAC, but the details of the coverage can be elusive. Before moving to a more in-depth discussion of LOAC, however, a look at other aspects of relevant international law is in order.

**CYBER ACTIVITIES OUTSIDE THE CONTEXT OF ARMED CONFLICT**

The most active area for international discussion relevant to cyber warfare is how cyberspace activities affect international relations and the possibility of resorting to cyber war or of cyber operations resulting in a war beginning. Of course, lawyers would prefer to confine the discussion to the legal issues. There is a body of law that governs the resort to war, but politics and relations between States are much more the issue with cyber warfare. The dance among States as they carry out trade, diplomatic relations, espionage, etc. is delicate. In the end, although the UN Charter provides the only lawful means of resorting to armed conflict, i.e., when sanctioned under Chapter 7 or in response to an armed attack, political and military leaders tend to talk less about the law in the area and more in terms of what constitutes an “act of war.”

The determination that something is an act of war expands the discussion beyond the law. It concerns the relative strength of the involved States, the domestic political situation, alliances, intelligence analysis, and more. These factors greatly outweigh legal considerations in the actual calculus of States. This is easy to see when hypothetically reversing parties in some actual cyber incidents. For example, if Iran had damaged a nuclear facility in the U.S. as the U.S. is said to have done in Iran with the Stuxnet virus, the victim’s public reaction to the operation would have been very different.

Similarly, if Estonia had engaged in cyber aggression against Russia equivalent to what Russia did to Estonia in 2007, it is likely Russia would

---


17. The article uses the term “war” here rather than the legal formulation “armed conflict” because the point is that war is partly a political decision, informed rather than controlled by law.


19. *See generally Gary D. Brown, Why Iran Didn’t Admit Stuxnet Was an Attack, 63 JOINT FORCES Q. 70 (2011). Israel was also implicated in the Stuxnet incident. For a full explanation of Stuxnet, see KIM ZETTER, COUNTDOWN TO ZERO DAY (2014).*
have responded aggressively in self-defense. Imagine Russia’s reaction to having large numbers of its government and banking websites offline for hours at a time over a period of several days. Taking a position consistent with the relative sizes of the States involved, however, Estonia determined the activity would be better handled as a criminal matter rather than a breach of international peace.

Ultimately, States’ judgments on whether they are the victims of an act of war that provides sufficient cause to engage in national self-defense is circumscribed by political reality and, while the law may inform the decision, it does not compel it.

To ensure clarity for the remainder of the paper, the following chart sets out a framework for the application of international law to cyber warfare. Although cyber means and methods are a part of warfare, war is also still caused and carried out by physical means. This article is meant to look at cyber-specific situations where there is little precedent and a great deal of ambiguity about how the law should operate.

---

20. ENEKEN TIKK, KADRI KASKA & LIIS VIHUL, INTERNATIONAL CYBER INCIDENTS LEGAL CONSIDERATIONS 18-25 (2010).

21. Id. at 25-26.

22. U.N. Charter art. 39. UN. UN authorization to use force under Chapter 7 is straightforward if relatively rare. States must decide on a regular basis whether hostile activities directed against them constitute an act of war, which they would characterize as armed attacks for purposes of public justification.
2017] INTERNATIONAL LAW APPLIES TO CYBER WARFARE!

The chart represents how the law applies to various effects. Below the ARMED CONFLICT band is peacetime (at least, non-armed conflict) operations. However, most of the time kinetic operations during peacetime instantly elevate the situation above the line. That is, they trigger armed conflict, although the conflict may be quite brief if the victim decides not to respond. It’s important to note that the applicable law is determined by the effects, not by the method. For example, if a cyber method causes a kinetic effect, it is treated no differently than if it were caused by a traditional kinetic means.

Operations below the line of armed conflict on the chart are not governed by the law of armed conflict. The bottom right box generally presents typical bellicose operations. If kinetic effects result (property destruction, injuries, or death), the situation may be pushed above the line to armed conflict—even if the kinetic effects are caused by cyber means or methods. The lower left box is the typical use of cyber techniques to annoy, harass, disrupt, and interfere outside of armed conflict. It’s unclear when, if ever, such activity alone can create a state of armed conflict.

Above the line, the top right box is typical warfare, involving destructive and injurious effects. Even if kinetic effects in the context of armed conflict are created with cyber means and methods, the application of LOAC is clear, and no different than if the effects were created with kinetic means. Destroying civilian structures or directly injuring civilians (perhaps through manipulating medical devices), when the places and people are the target of the cyber attack, is unlawful, but there is little mystery there. Some problems, like electrical power, are trickier, but the rules that have worked for bomb dropping should work equally well for cyber techniques. If anything, LOAC should operate to encourage cyber over kinetic operations because it’s likely the civilian impact will be less when a system isn’t destroyed as it is with kinetic options, but rather is rendered non-usable with cyber means and can be turned on again after the conflict. If LOAC operates to permit a broader range of cyber activities in war, civilian death and destruction will tend to be diminished. Rather than having only force to achieve national security goals during armed conflict, States could also have effective, lawful cyber options.

The unique aspect of the use of cyber means and methods in warfare is represented in the top left box. In the context of armed conflict, cyber

23. See discussion on different types of armed conflicts infra pages 363-65.
techniques that do not result in kinetic effects create new issues for LOAC. It is this specific category of activities that this paper addresses.

GETTING TO CYBER WAR

LOAC, which is sometimes referred to as the law of war or international humanitarian law (IHL), is best defined as “...the controlling body of law with respect to the conduct of hostilities and the protection of war victims.”

As might be obvious from the name, LOAC applies during an armed conflict. LOAC is comprised of a specific set of principles that apply in a distinct situation that involves violence, destruction, injury, and death. It applies only when States or armed groups have broken the peace. At least one (and often more) of the parties involved in the armed conflict have already shown disdain for legal constraints on behavior by their resort to violence in the first place. Still, the body of law applies to all parties involved in an armed conflict, regardless of whether the conflict is just or unjust, and no matter who started it. In the context of this violent situation, certain basic rules have been found to apply: the principles of distinction, military necessity, proportionality, and humanity.

LOAC is critically important for regulating conduct in warfare. It limits the use of inhumane weapons, prohibits the targeting of civilians and civilian property, and guards the wounded and captured, among other things. However, it is limited in application to armed conflict. Relevant to this discussion, that means LOAC applies to cyber warfare, but not to cyber activities outside the context of armed conflict. As noted previously, those lesser activities are not ungoverned by law, but they are untouched by this particular body of law. Just as the body of Virginia traffic laws, while perfectly valid and important, does not govern driving in Canada, LOAC has no authority to regulate cyber conduct outside the context of armed conflict.


26. LAW OF WAR MANUAL, supra note 25.

27. See generally id. It also serves the vital function of governing the treatment of detainees, but that aspect of LOAC is beyond the scope of this article.

28. U.S. DEP’T OF DEF., DIR. 2311.01E, DO D LAW OF WAR PROGRAM para 4.1 (May 9, 2006). DoD notes its policy that DoD members “comply with the law of war during all armed conflicts, however such conflicts are characterized, and in all other military operations.” While this provision is the subject of discussion in other contexts, it does not change the outcome with regard to cyber activities. The primary principles would fail to attach to non-attacks in non-armed conflict situations, just as in armed conflict.
Although it is straightforward that LOAC applies during armed conflict, it is not always easy to determine the existence of an armed conflict. Not every occurrence of violence is “armed conflict.” Attempts to define precisely armed conflict have been unsatisfactory, such as this one from Uppsala Universitet: “An armed conflict is a contested incompatibility which concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths.” 29 Although it was a noble effort, it is apparent that a specific definition like this raises as many questions as it answers. Why must a government or territory be involved? Why not 24 deaths? What if there are thousands of injuries but no deaths? How can massive property damage and destruction not result in a state of armed conflict?

Armed conflict can be subdivided in various ways. Uppsala Universitet’s creative definition lumps it all together, but international law requires armed conflict to be divided into two parts. That is because LOAC is comprised of two similar but legally distinct sets of rules. One governs international armed conflict (IAC) and the other non-international armed conflict (NIAC). 30

The first is simple to define. An international armed conflict is a resort to armed force between States. 31 This is the classic case of warfare—the government of one territorial State waging war against the government of another territorial State. There is generally not considered to be any specific threshold of death or destruction. A single shot fired in anger between two States results in a state of IAC. 32

It is challenging to conceive of a cyber-only IAC that did not include a large-scale kinetic effect. Cyber techniques that are used to cause fires, flooding, or mass transit accidents, for example, seem sufficient to meet the single shot threshold, just as the same event caused by kinetic actions (bombs, saboteurs, and assassins) would. To start an armed conflict without kinetic effects would be breaking new ground, but one focused on UN Charter law, i.e., the jus ad bellum rather than LOAC.

30. 31st International Conference of the Red Cross and Red Crescent, Nov. 28-Dec. 1, 2011, Geneva, Switz., International Humanitarian Law and the Challenges of Contemporary Armed Conflicts, 31IC/11/5.1.2, at 3 (Oct. 2011). In the LOAC discussion, this paper considers only principles common to both IAC and NIAC law.
Turning to non-international armed conflict, it is unhelpfully defined as “armed conflict not of an international character.”\(^{33}\) In addition to being an armed contest between entities that include at least one non-State, to be classified as armed conflict non-international clashes require some level of organization on the part of the non-State group, as well as some level of intensity in the violence to qualify as armed conflicts.

Various tribunals have considered the level of intensity required for armed conflict through evaluating a number of factors, some of which are particularly ill-suited to address groups of cyber actors, such as displacement of people due to the conflict and the number and type of weapons used.\(^{34}\) Other factors aren’t much better; they include the gravity of attacks and their recurrence, the expansion in territory, and duration of violence.\(^{35}\)

If anything, the analysis gets harder when considering the organization of cyber actors. The law requires the non-state actors to be armed enough so that they have the capacity to mount attacks.\(^{36}\) Organizational factors assessed include whether the group has internal regulations; whether it can issue orders and coordinate attacks effectively; the establishment of disciplinary rules and enforcement mechanisms; the ability to recruit members; and the use of uniforms.\(^{37}\)

It seems unlikely these factors would be present sufficiently for there to be a cyber-only NIAC. Much of the nefarious activity on the internet is undertaken by hacktivists and loose collections of actors such as Anonymous, the Chaos Computer Club, and LulzSec, which hardly qualify for the moniker “group,” much less “organized group.”\(^{38}\)

\(^{33}\) Geneva Convention (III) Relative to the Treatment of Prisoners of War, art.3, Aug. 12, 1949, 6 U.S.T. 3316, 75 U.N.T.S. 135. To make matters even more confusing, the ICRC recognizes six different types of NIAC, and takes note of a seventh type recognized by some. This level of discussion is well beyond the scope of this paper, but is set out in 31st International Conference of the Red Cross and Red Crescent, Nov. 28-Dec. 1, 2011, Geneva, Switz., International Humanitarian Law and the Challenges of Contemporary Armed Conflicts, 31IC/11/5.1.2, at 9-10 (Oct. 2011).


\(^{36}\) See LAW OF WAR MANUAL, supra note 25, at 74-76.

\(^{37}\) Limaj, Case No. IT-03-66-T at ¶ 94-129

Whether a state of armed conflict is triggered by cyber activities and effects or, as seems more likely, by kinetic events, once a state of armed conflict exists, combatant activities are governed by LOAC. The principles that make up the essential body of LOAC are noted above. This paper will focus on two of them, distinction and proportionality.

**CONDUCT OF HOSTILITIES WITH CYBER MEANS: A LIMITED ROLE FOR LOAC**

“Distinction requires parties to a conflict to discriminate in conducting attacks against the enemy. . . . parties may not make the civilian population and other protected persons and objects the object of attack.”

What is referred to as the proportionality rule would prohibit “an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated.”

Note the word “attack” in both principles. Proportionality and distinction only apply in cases of “attack” in the context of armed conflict. So, as simple as seems, the principles cannot be applied properly without a definition of attack.

The obvious question to be answered, then, is what exactly is a cyber attack? The definition of attack in international law is “acts of violence against the adversary.” Cyber operations that lack direct physical effects are not violent and so cannot be classified as attacks. The Tallinn Manual’s definition for cyber attack is “a cyber operation . . . that is reasonably

---

40. *LAW OF WAR PROGRAM, supra* note 23, at para 2.5.2. The DOD’s definition, rather than an international definition, is used here for reasons discussed *infra* at pages 371-72.
41. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), art. 51(5)(b), June 8, 1977, 1125 U.N.T.S. 3; *The U.S. is not a party to* Add’l Protocol I *but considers large parts of it, including this provision, to be accurate statements of customary international law.*
expected to cause injury or death to persons or damage or destruction to objects."

So, even though it is clear LOAC applies to cyber activities inside an armed conflict, its relevance is limited. The two most important principles of LOAC, distinction and proportionality, both attach on attacks. That is to say, activities that are something other than attacks do not trigger application of the principles.  

Cyber attack must be distinguished from cyber disruption. The term “cyber disruption” is used here to refer to cyber only operations that cause inconvenience, even extreme inconvenience, but no direct injury or death, and no destruction of property. There have been many examples of these kinds of effects caused by computer malfunctions. Considering how such events would be characterized if they had been intentionally caused may help illustrate why they should not be categorized as attacks.

In 2016, both Delta Airlines and Southwest Airlines suffered major disruptions of service when computer systems malfunctioned. Both airlines were forced to ground hundreds of flights, losing millions of dollars in revenue. There was no injury or damage, but major financial losses. If the computer systems were destroyed by bombing in an armed conflict, they would be considered attacks, of course. What if instead the problems were caused, during an armed conflict, by hiring all the competent computer operators away from the airlines, or by stealthily changing the cipher lock combinations on the doors to the computer facilities? Both of these could result in the same disruptions to the computer networks, but would not be classified as attacks. Neither should the same result caused by a non-

43. See TALLINN MANUAL 2.0, supra note 4, at 415.
44. See generally, MARCO ROSCINI, CYBER OPERATIONS AND THE USE OF FORCE IN INTERNATIONAL LAW 239-40 (2014). The author there does, however, take a contrary view on the application of Add'1, Protocol I, Art. 57, Precautions in Attack.
destructive cyber operation be defined as an attack, despite its deleterious effect on the civilian population.

Similar actions could be designed to aid in a military campaign, without the actions themselves being attacks. Compromised networks controlling transportation systems, such as railroad switching and air traffic control, might be manipulated so they become buggy and unreliable. This would likely force the authorities to halt or curtail traffic to prevent accidents. Similarly, military operators might cause networked traffic lights in a major city like Los Angeles to work only sporadically, or set them all to red, snarling traffic. Software controlling ship traffic in Houston, New York, Long Beach, and other major seaports could be caused to crash, tangling the unloading of ships and causing shortages of items across the country. Penetrated financial networks might be manipulated to delete, modify, or transfer balances in individual bank accounts. None of these activities would cross the threshold of an attack, just as the same results caused with non-destructive physical means would fall short of an attack. For example, if an extension cord were unplugged causing equipment to go to a battery backup with limited functionality, or if phone calls were placed to harbormasters to trick them into sending ships away without unloading, there would be no military attack, but the result of the activity would be disruptive to an enemy in its war efforts.48

The unique aspects of cyber warfare—speed, ubiquity, and lack of geographic constraints—have brought this issue to the forefront. The principles of LOAC were agreed to and have been practiced in a pre-cyber warfare world. The introduction of cyber capabilities does not change the basic principles of LOAC, which continue to require attacks for the principles to apply.

The mischief that can be caused with cyber disruption, while still falling below in bello attack, illustrates how the time-honored LOAC principles, while very useful in a traditional kinetic situation, fail to provide the same level of protection for civilians who might be victims in cyber war. Experts in the field take for granted that the principles will be relevant in every situation that might be averse to the interests of civilians caught in armed conflict. Actually reading the words of the concepts reveals that this is not so.

Observing the gap cyber warfare techniques appear to have opened in civilian protections during armed conflict, some international law experts have suggested the definition of attack, with regard to cyber, should be

---

48. Of course, any or all of these activities might be criminalized under domestic law as sabotage or some other crime. The point here is that such actions are not limited by the LOAC principles of distinction and proportionality.
expanded to include loss of functionality. This would not merely be an application of existing law to a new method of warfare. This would be a redefinition of a term of art beyond anything it has previously been found to mean.

The Tallinn Manual breaks the functionality issue into three basic scenarios. A cyber operation can physically damage a component of a computer system, can cause it to cease functioning until the operating system is reinstated, or can cause it to cease functioning by deleting or interfering with data on the system (e.g., the targeted computer still functions as a computer, but isn’t functional as a communications node because the communications program has been deleted).

The ICRC would like to see it expanded this way. “[I]t is clear that the damage to be taken into account comprises not only physical damage, but also the loss of functionality of civilian infrastructure even in the absence of physical damage.” More specifically, the ICRC argues:

[T]he fact that a cyber operation does not lead to the destruction of an attacked object is also irrelevant. Pursuant to article 52(2) of Additional Protocol I, only objects that make an effective contribution to military action and whose total or partial destruction, capture or neutralization offers a definite military advantage, may be attacked. By referring not only to destruction or capture of the object but also to its neutralization the definition implies that it is immaterial whether an object is disabled through destruction or in any other way.

The preeminent legal scholar in the field of cyber warfare offers a somewhat different view. “I am nevertheless now persuaded by the foundational premise of the restrictive approach, i.e., that the notion of cyber attacks cannot be limited to injurious or physically destructive cyber operations . . . My own view is that a system has lost functionality when it is no longer able to perform its intended function without some repair. This would include reloading the operating system or any software essential to its operation, but would not include replacing data that was merely stored on the system.” Although this is a reasoned position, it is still an expansion of existing law.

49. See TALLINN MANUAL 2.0, supra note 4, at Rule 92, para 10, at 417.
50. Id. at Rule 92, paras. 10-12, at 417-19.
51. Id.
52. Droege, supra note 15, at 571.
53. Id.
2017] INTERNATIONAL LAW APPLIES TO CYBER WARFARE! 369

Under the terms of the Tallinn Manual taxonomy discussed above, the ICRC advocates for the broadest definition, which would define any cyber event causing a loss in functionality as an attack. Professor Schmitt argues for the middle option. Both of these approaches create issues under current law, although the ICRC approach is more problematic.

CONSEQUENCES OF APPLYING A FUNCTIONALITY STANDARD

The definition of function is “the kind of action or activity proper to a person, thing, or institution; the purpose for which something is designed or exists.” For example, the primary function of cell phones is to act as communications devices, the primary function of a bridge might be to provide a path across a river, and the primary function of car is to transport a person between two places. In kinetic operations, the method of preventing the function of these items might be to bomb cell towers, blow up the bridge, and crush the car by driving over it with a tank. Clearly, all these options are attacks, and the principles of LOAC would apply. The attacks would have to be discriminate, necessary, proportionate, etc. to pass legal muster.

On the other hand, what if a military commander decided to approach the problem differently, by buying up all the cell phone service provider’s bandwidth to prevent the operation of local cells phones, by parking military vehicles on the bridge to prevent the passage of civilian traffic, and by taking the car keys from the car. All of these techniques prevent the function of the civilian objects. However, none of them looks like an attack, and none would be analyzed as such under the principles of the law of war. Why should cyber activity be treated differently? If a denial of service against the cell phone service provider, a hack that raises a drawbridge, and reprogramming car key fob to render it inoperable all can have the same effect as non-destructive kinetic events, why should they be analyzed differently?

If the ICRC approach were the law, other fairly innocuous actions in the physical realm would also be considered “attacks,” because they similarly interfere with the functionality of an object without damaging it. The

---

Professor Schmitt agrees to some extent with ICRC’s argument, he rejects its basis for the conclusion.


56. Each of the examples has other functions, as well, but it cannot be the case that every possible function of a device must operate lest the activity that ended that function be considered an attack. A cell phone may function as a paperweight, but picking it up from a stack of papers does not damage its functionality.

57. That’s not to say the actions wouldn’t be analyzed for policy and strategic reasons. Actions in and out of warfare are often avoided if they are “lawful but awful.”
expansion in the law wouldn’t be limited to cyber activities, it would also extend to kinetic activities with similar effects. Activities in wartime such as hiring civilian truck drivers, using roadways, or letting the air out of tires all reduce the functionality of civilian trucks, but none of these activities is an attack, and there would be no consideration given as to whether hiring civilian drivers violates the proportionality principle or whether driving on a roadway violates the principle of distinction, for example. The ICRC approach would appear to render all of them attacks, which is simply not the law.

Professor Schmitt’s approach specifies that some sort of repair would be required before a loss of functionality would equal an attack. This is closer to what the law requires, but still appears to expand it from its current state. For example, draining a battery necessitates recharging the battery, a type of repair. Would turning on a truck’s lights, which might result in draining the battery, constitute an attack in bello? If a cyber attack could remotely drain a system battery by causing a screen to stay on at full brightness, for example, would that be an attack? It is difficult to think of a good kinetic analog to recharging system software, but perhaps it’s akin to stealing an instruction book so that equipment can’t be operated. Would such a theft be considered an attack? Referring to the previous paragraph, is adding air to a deflated tire a repair?

If these examples seem absurd, it is because they are. LOAC was designed to provide broad legal coverage of destructive wartime activities to protect civilians from death, injury, and property destruction, not to prohibit disruptions or inconveniences. As discussed earlier, LOAC should encourage non-destructive, non-lethal cyber activity in order to hasten a return to normalcy post bellum.

The role of cyber operations in national security is important, and growing in importance, but once an armed conflict begins, generally cyber warfare fades to the background in the white heat of kinetic battle. Cyber operations are in support, providing options to help degrade the adversary’s ability to counter actions. Now and for the foreseeable future they will be the smallest concern when weighed against death, injury, and destruction.

58. Schmitt, supra note 54, at 203.
THE FUTURE

LOAC is a real body of law. Without new treaties or established custom, it won’t change. In the area of cyber warfare, it is generally considered that a treaty is unlikely to the point of impossible because of international disagreement over basic principles such as free speech and privacy. Cyber warfare also presents an especially difficult case study for the development of norms, as States have to date kept their cyber operations concealed. That means there has been no progress toward developing common practice.

As the two common methods of changing the law are unlikely to result in change with regard to cyber warfare, any unique issues that arise must be dealt with by interpreting existing law, rather than writing new. The problem with creatively trying to adjust LOAC to fit the situation is that it could make it more difficult to apply and thus less effective for kinetic warfare.

Until there is a change in the political landscape, which could happen very quickly if there is a catastrophic cyber event, there will remain a gap in the protection of civilians from inconvenience in warfare. Despite this gap, the risk nondestructive cyberspace operations pose to civilians is outweighed by the risk of damaging LOAC’s application in traditional armed conflict.

RED HERRINGS: EXPANDED DISTINCTION, THE MARTENS CLAUSE, AND TELEOLOGY

The discussion above did not use the most common international statement of the principle of distinction, which is found at Add’l. Protocol I, Art. 48: “Parties to the conflict shall at all times distinguish between the civilian population and combatants and between civilian objects and military objectives and accordingly shall direct their operations only against military objectives.” This definition of the principle is misleading, in that it uses the phrase “operations” rather than attacks when all the examples Add’l. Protocol

59. States develop new international law through their custom and practice, followed by eventually believing they are legally bound to act in certain ways because it is the practice. This can be aggravating for scholars, who would prefer open discussion and full transparency regarding the development of international law. This issue is fully discussed in Michael N. Schmitt & Sean Watts, State Opinio Juris and International Humanitarian Law Pluralism, 91 INT’L L. STUD. 171 (2015), http://stockton.usnwc.edu/ils/vol91/iss1/6.


I provides are, in fact, attacks. Some scholars argue that it should be read to cover actions that would not qualify as attacks. Actions suggested for coverage are “all the movements and activities carried out by armed forces related to hostilities.”

Despite the complex academic rationale in favor of this interpretation, two compelling arguments run counter. First, if the Add’l. Protocol I drafters intended the terminology to apply to situations broader than attacks, they would have offered examples of what those situations might be. Instead, all the examples provided concern attacks, which is consistent with the way the Protocol defines proportionality.

Second, and even more compelling, the broad interpretation is simply not how States have applied this provision in practice. When reviewing military activities in the context of operations, logistics, communications, and other support activities are not reviewed for compliance with the principle of distinction. The effect of such operations on civilians may be reviewed for practical or policy reasons, such as compliance with rules of engagement, but no legal review is authorized to determine if the act of landing cargo planes on a civilian runway or using a civilian radio frequency to transmit military supply communications violates the principle of distinction, even though both fail to distinguish between civilian and military objects in the same way analogous cyber operations would.

For these reasons, the U.S. statement of the principle more accurately reflects the reality of the law than Add’l. Protocol I, and so the U.S. statement was used in the main body of this article.

It has been similarly argued that another provision of Add’l. Protocol I might be relevant here. Supplemental to, or in addition to (it is unclear which), the principle of distinction is Add’l. Protocol I, Art. 51, which provides civilians “general protection against dangers arising from military operations.” This provision has been a topic of discussion among international legal experts, but ultimately fails to change the conclusion that civilians and civilian objects may be targeted by cyber operations that fall below the level of an attack.

62. Id. at 600.
63. INT’L COMM. OF THE RED CROSS, supra note 61.
64. A summary of this position may be found in HEATHER HARRISON DINNISS, CYBER WARFARE AND THE LAWS OF WAR 196-202 (2012).
65. It certainly more accurately reflects customary law, and the section on Add’l. Protocol I, Art. 51 that follows, suggests it better represents the law of Art. 48, as well.
67. Id. at 613.
The provision does not alter the conclusion about cyber warfare because, based on the examples given in the rest of Art. 51, it was not included to add restrictions to non-attack military operations. Rather, it was written to emphasize that certain types of egregious attacks on civilians are prohibited. The lack of clarity in the wording of the provision is noted by eminent international jurist Yoram Dinstein. “It is not clear what dangers arising from military operations—other than attacks—the drafters of AP/I had in mind.” The ICRC Commentary asserts the language should apply to all the activities armed forces carry out pursuant to armed conflict. Although the sentiment is laudable, it reflects neither the language of the provision nor State practice in this regard. Troop movements, billeting, and security measures often inconvenience the civilian population, but States do not apply the principles of warfare to make decisions about these operations. For example, military leaders do not attempt to determine if driving a convoy on a road is “indiscriminate” just because the road is a civilian object whose functionality will be impaired while it is being used for military purposes.

After discussing possible interpretations of the Add’l. Protocol I, Art. 51 language, Professor Dinstein ultimately concludes the provision must be read to apply only to attacks, expressing concern that it leaves open the door for sub-attack cyber operations to be directed against the civilian population.

Neither Art. 48 nor Art. 51 of Add’l. Protocol I, as noted above, moderate adverse cyber warfare effects on civilians satisfactorily. When something appears to fall outside the jurisdiction of LOAC, especially if there is ambiguity in the situation, scholars sometimes turn to the Martens Clause. The Martens Clause, from the Preamble to the 1899 Hague Convention, is

68. Id. at 613.
69. “To give effect to this protection, the following rules . . . shall be observed.” INT’L COMM. OF THE RED CROSS, COMMENTARY ON THE ADDITIONAL PROTOCOLS OF 8 JUNE 1977 TO THE GENEVA CONVENTIONS OF 12 AUGUST 1949, supra note 62, at 613.
70. Id. at Art. 51(2) and (4) at 613.
73. The issue is discussed in the context of battlefield biometrics, but is not brought to a definitive conclusion in Alison Mitchell, Distinguishing Friend from Foe: Law and Policy in the Age of Battlefield Biometrics, 50 CAN. Y.B. INT’L L. 289, 309-10 (2012).
74. DINSTEIN, supra note 71.
75. There is a nearly identical argument with regard to Art. 57. See DINNIS, supra note 49, at 200-02.
named for its author, Russian Professor von Martens. “Until a more complete code of the laws of war is issued, the High Contracting Parties think it right to declare that in cases not included in the Regulations adopted by them, populations and belligerents remain under the protection and empire of the principles of international law, as they result from the usages established between civilized nations, from the laws of humanity and the requirements of the public conscience.”

Although it is a valuable tool in some contexts, with regard to cyber warfare, the Martens Clause adds nothing to the mix. LOAC already applies, so the Clause is unnecessary to ensure legal coverage. The big question is exactly how the law applies to cyber operations, and the language of Martens, being quite general, adds no clarity to that.

Finally, it may be argued that failing to apply LOAC principles to cyber disruption targeted at civilians violates the purpose of LOAC. After all, ICRC defines IHL as “a set of rules which seek, for humanitarian reasons, to limit the effects of armed conflict. It protects persons who are not or are no longer participating in the hostilities and restricts the means and methods of warfare.” This definition, however, is overly restrictive in that it reflects only one rationale for LOAC, and notes only its limiting function, which is why this paper used another definition for its analysis.

From the earliest attempts to develop a formal body of law to govern warfare there was a recognition that implementing general protective rules would facilitate a return to peace. A practical body of wartime law facilitating a return to peace is more likely to motivate States to comply than would a protective code created without a recognition of the unfortunate reality of war. States desire peace not only because it benefits civilians, but also because it generally serves the security interests of States.


79. Byron D. Green, Bridging the Gap that Exists for War Crimes of Perfidy, 2010-AUG. ARMY L. 45, 3.2 (2010) (explaining that the purpose of the LOAC is to “humanize warfare to the maximum extent possible.”).


81. See McLeod, supra note 20 (explaining that the LOAC is “the controlling body of law with respect to the conduct of hostilities and the protection of war victims”).

A restrictive definition of LOAC, which would result in the application of principles that would limit cyber disruption during armed conflict, does not reflect the law or State practice. Further, although it may seem to reflect a kinder legal regime, in the long run it might be less humane. If States are permitted to employ a broad range of disruptive cyber options, even against civilian objects, in many cases it could offer a rationale alternative to the collateral damage and casualties likely to result from lawful kinetic attacks.

As John Fabian Witt notes in his book on the law of war, Lincoln’s Code, the original U.S. law of war code was “not just a humanitarian shield . . . [i]t was also a sword of justice.” If States are permitted to employ a broad range of disruptive cyber options, even against civilian objects, in many cases it could offer a rationale alternative to the collateral damage and casualties likely to result from lawful kinetic attacks. LOAC’s remit is broader, providing a legal framework for warfare while constraining its effects for the benefit of civilians and, ultimately, the parties to the armed conflict.

CONCLUSION

There are many challenges in applying international law to cyber activities. The least of these are concerns that arise during armed conflict. The critical issues are those surrounding cyber operations outside existing armed conflict. Uncertainty outside the context of existing armed conflict could cause international miscalculation, breaking of the peace, and a rapid escalation of aggression. There is work to be done to increase confidence among States, develop norms of appropriate cyber behavior, and to provide privacy guarantees to citizens, among other things. None of these issues is within the ambit of LOAC.

As discussed previously, LOAC has always covered cyber warfare. The issue has been concern over how it does so. LOAC, being a practical body of law, recognizes civilians will be adversely affected by war. Inconvenience and bother are the smallest concerns in warfare. LOAC addresses death and devastation, including that caused by cyber warfare. It just does not prohibit nonviolent actions that would cause civilians to drive

84. See Huntley, supra note 79, at 2 (explaining that while the LOAC may not be completely effective when applied to cyber warfare, the LOAC, nevertheless, applies to cyber warfare).
85. See id. at 2-3 (discussing concern over the ineffectiveness of LOAC to deter cyber warfare).
the long way around to work, to lose cable TV, to be deprived of their favorite soda . . . nor does it protect them from being cut off from social media. In other words, what have to date been the most common uses of cyber capabilities operate below the level at which LOAC would restrict them. No attack means no proportionality or distinction analysis. When cyber attacks cause kinetic effects, by damaging a piece of industrial equipment, for example, analyzing the damage is the same regardless of whether it was caused by a saboteur, air-delivered ordnance, an artillery shell, or by a cyber attack. No cyber-specific analysis is required, or helpful.

The functionality gap discussed here has caused consternation in the international legal community, with some members fearing civilians might suffer as a result. So far, even without LOAC governing disruptive cyber activities, civilians have not suffered greatly as a result of cyber warfare. It may be that States have determined as a matter of policy not to carry out such activities, that they think wartime cyber disruption of civilians would not advance their strategic interests, or that warring States simply lack the capacity for an effective disruption campaign. Particularly if one of the first two is the explanation, there are grounds for hope that customary law might develop that would offer the civilian community more formal protection.

Although LOAC has little to offer in controlling unique aspects of cyber operations in the context of armed conflict, it continues to play a vital and irreplaceable role in regulating kinetic operations in warfare. LOAC is fragile by nature because it is designed to affect behavior between warring parties, who do not like each other to start with, and may perceive major advantages in violating the law. The body of the law of armed conflict has proven resilient through the years, but it is important to avoid stretching it to the breaking point. Broadening definitions for the purpose of achieving more direct coverage of cyber activities during armed conflict risks undermining LOAC and losing its application to kinetic operations, where it is a critical and historically proven stay on the lethal and destructive activities of States.

It may be that permitting non-attack cyber operations to target civilian objects in armed conflict is not the best answer. In that case, States must

86. See, e.g., Ariana L. Johnson, Cybersecurity for Fin. Institutions: The Integral Role of Info. Sharing in Cyber Attack Mitigation, 20 N.C. BANKING INST. 277, 303-04 (2016) (discussing how cyber warfare can cause physical harm and economic harm, including identity theft).

87. See, e.g., id. at 277-81 (explaining that while banks are consistent victims of cyber warfare, banks constantly manage to recover and build new technologies to fight cyber warfare; thus, they avoid long-term financial damage or collapse).

88. See, e.g., Ryan J. Vogel, Drone Warfare & the L. of Armed Conflict, 39 DENV. J. INT’L & POL’Y 101, 137 (2010) (discussing how the LOAC has been successful in regulating sophisticated technology, like drones).
determine whether non-attack cyber effects in armed conflict are best governed by the aggressively creative interpretation of LOAC critiqued here, or rather by some other body of law or norms. Until States, the masters of international law, decide to change the law, it will remain what it is—adaptable and effective in its rules protecting civilians from the effects of attacks in warfare, and ambivalent about activities that fall below that level.
“Cybervandalism” or “Digital act of war”? America’s muddled approach to cyber incidents won’t deter more crises

By Charlie Dunlap, J.D. - 30 October 2016

If experts say a malicious cyber code has “similar effects” to a “physical bomb,” and that code actually causes “a stunning breach of global internet stability,” is it really accurate to call that event merely an instance of “cybervandalism”?

Moreover, can you really expect to deter state and non-state actors from employing such code and similarly hostile cyber methodologies if all they think they are risking is being labeled as a cyber “vandal” subject only to law enforcement measures? Or might they act differently if it was made clear to them that such activity is considered an “armed attack” against the United States and that they are in jeopardy of being on the receiving end of a forceful, law-of-war response by the most powerful military on the planet?

Of course, if something really is just vandalism, the law enforcement paradigm with its very limited response options could suffice. But when a malevolent cyber activity endangers the reliability of the internet in a world heavily-dependent on a secure cyberspace, it isn’t just vandalism. Rather, it is a national and international security threat that ought to be characterized and treated as such. Unfortunately, the U.S.’s current approach is too inscrutable and even contradictory to send an effective deterrence message to potential cyber actors. This needs to change.

“A stunning breach of global internet stability.”

If you need an example of the U.S.’s cyber deterrence problem, consider the massive disruption of the Internet that took place on October 21st, 2016. Twitter, Paypal, Spotify and other many other popular websites were virtually shut down when Dyn, a domain name system (DNS) provider that functions as a “switchboard” for an enormous amount of internet traffic, was the victim of a huge distributed denial-of-service (DDoS) attack. How serious was the Dyn incident? William Turton, writing on Gizmodo, said that half the internet was shut down, and Reuters characterized the cyber crisis as “a stunning breach of global internet stability.”

What made this cyber incident especially worrisome is that experts believe the “attackers apparently used tens of thousands of hacked internet of things devices — household appliances such as digital video recorders, security cameras, and internet routers — to generate a massive amount of digital traffic” that jammed the system and ground it to a halt several times.

Although Web functionality was more or less reconstituted by the end of the day, the Dyn attack may signal things to come. A retired intelligence officer ominously suggested that it may have been a probing attack, that is, one designed to enable a hostile actor to “eventually launch a devastating, Pearl Harbor-type cyber attack.” Even before the most recent incident, Bruce Schneier said the “precisely
calibrated [cyber] attacks” of recent months “feels like a nation's military cyber-command trying to calibrate its weaponry in the case of cyberwar.”

Nation-states are not the only ones who need to be deterred from nefarious internet activity as non-state actors can also cause serious disruption. For example, regarding the Dyn incident, Mr. James Clapper, the Director of National Intelligence, indicates that “preliminarily” a non-state actor may be responsible. Regardless, the vulnerability to a range of hostile actors is painfully evident: the devices exploited in this event – which are made with some parts “coming from Chinese suppliers [and] have weak or no password protections” – are extremely common. Intel Corporation predicts the world will have 200 billion of them by 2020, so it’s unlikely that we’ve seen the last of these cyber emergencies.

Cyber vandalism or digital acts of war?

Surveying the wide-ranging impact of last week’s Web calamity, an analyst observed that, “[e]ven though [the malware involved in the attack is] not a physical bomb, it has some similar effects.” The question then is this: does the U.S. consider this unprecedentedly severe incident, involving as it does a cyber capability that has “similar effects” to a “physical bomb,” to be a digital act of war?

Evidently not. Even though the facts of the massive shutdown would seem to equate the incident with a traditional kinetic attack, NBC news reports a senior U.S. intelligence official as rather dismissively classifying the incident as just “a classic case of internet vandalism.”

Does the official’s characterization conform to what the U.S. has said previously about the legal status of certain cyber events? Of course, it’s important to understand that “act of war” is a political term, not one of international law. In the post-UN Charter era, the “act of war” idiom is at odds with the underlying thrust of the Charter and especially Article 2(4). It demands that: “[a]ll members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations.” In essence, “war” as historically understood, is all but illegal; disputes are to be resolved by peaceful means.

There are, however, narrow exceptions to the prohibitions against the use of force. Force is allowed when the Security Council authorizes it under Article 42 of the Charter. Additionally, a nation may employ force in self-defense when it suffers what Article 51 describes as an “armed attack.” (Most nations also believe Article 51 also incorporates an inherent right to act in anticipatory self-defense when an armed attack against them is imminent.)

The law does not define exactly what kind of form of forceful response a country may take in a legitimate act of self-defense except to say it must be necessary and proportional. Nor does it limit a self-defense response in a cyber-situation to an “in-kind” response. The U.S., for example, says a self-defense response to a cyber-attack might include the use of traditional kinetic force involving conventional military weapons.

So when does the U.S. consider itself to have suffered an “armed attack” in the cyber context so as to trigger a right to self-defense under Article 51? Despite the enormous dimensions of the Dyn onslaught, the official’s claim that it is simply “cyber vandalism” (as opposed to any sort of “attack”) seems to suggest that the U.S. doesn’t consider it serious enough to permit a self-defense response within the meaning of Charter. This characterization is rather ironic as the U.S. has previously expressed a rather aggressive stance regarding what sort of cyber incidents could authorize forceful acts in self-defense.

In a seminal 2012 speech, the then-Legal Advisor to the State Department Harold Koh staked out the U.S. position. Initially, he affirmed that “established jus ad bellum rules do apply to uses of force in cyberspace.” (Jus ad bellum is that “branch of law that defines the legitimate reasons a state may engage in war and focuses on certain criteria that render a war just.”) He went on to explain that cyber activities that “proximately result in death, injury, or significant [physical] destruction would likely be viewed as a use of force.”
Pointedly, Koh also said that “if the physical consequences of a cyber attack work the kind of physical damage that dropping a bomb or firing a missile would, that cyber attack should equally be considered a use of force.” (Emphasis added.) In light of the claim that last week’s incident has “some similar effects” to a “physical bomb,” was the Koh threshold met? Or does the absence of “death, injury, or significant destruction” make it fall short in the U.S.’s view?

It isn’t clear. Koh makes it hard to determine because he said that the U.S. “has for a long time taken the position that the inherent right of self-defense potentially applies against any illegal use of force” adding that “[in the U.S.’s] view, there is no threshold for a use of deadly force to qualify as an ‘armed attack’ that may warrant a forcible response.” In other words, from the U.S. perspective, there is no difference between “force” as used in Article 2 of the UN Charter and “armed attack” as used in Article 51.

This interpretation of the law is, internationally, a distinctly minority view, as Professor Michael Schmitt and other cyberlaw experts have noted. It creates a complication because most interpretations of international law find that there are actions which might constitute “force” under Article 2, but not involve the kind of proximate “death, injury, or significant destruction” typically associated with an “armed attack.” Citing Nicaragua v. U.S., Schmitt provides an illustration with obvious implications for the U.S. position on cyber uses of force:

[T]he International Court of Justice held that although merely funding guerrillas who were conducting hostilities against another State did not reach the use of force threshold, arming and training them did. The holding suggests that an act need not have immediate physical consequences to comprise a use of force. (Emphasis added.)

When the Nicaragua holding is juxtaposed with Koh’s assertion that “force” and “armed attack” are conterminous, it seems that the U.S. should consider a grave cyber event like the Dyn attack as the legal equivalent to an “armed attack” even if it did not produce “death, injury, or significant destruction.” After all, if the U.S. position is that any use of force is enough to justify an Article 51 response, disrupting half the global internet with a methodology with effects similar to a “physical bomb” would certainly seem to be at least as significant as arming and training guerrillas in a single country.

To consider an incident as severe as the Dyn case as sufficient to put the perpetrators at risk of a forceful self-defense response not only would conform to the existing U.S. interpretation, but also could signal a norm evolution consonant with what some in the international community are already coming to realize about cyber attacks and Article 51 of the Charter.

The 2013 Tallinn Manual, which many consider to be the leading treatise on the international law applicable to cyberwar, does find that “force” as used in Article 2 (4) is different from the arguably more egregious “armed attack” as set out in Article 51. At the same time, however, its included commentary reports that the group of experts who drafted the Tallinn Manual found the law was “unsettled” as to whether “actions that do not result in injury, death, damage or destruction, but which otherwise have extensive negative effects” could amount to an armed attack.

In fact, we may be seeing a shift towards broader acceptance of the idea that cyber incidents with widespread adverse effects are enough to trigger an Article 51 response, even without any physical injuries or damage. In 2015, two years after the issuance of the Tallinn Manual, Professor Schmitt, who was the project’s director, agreed that if a cyber operation shut down the national economy without death or destruction, it would nevertheless “probably” meet the more demanding “armed attack” threshold.

In addition, UCLA’s Professor Kristen Eichensehr noted the conundrum that “cyber weapons create the possibility of actions that cause severe harm to the victim, but nevertheless do not result in physical damage or injury to persons.” Consequently, she predicted, it is “possible that over time a cyber specific definition of armed attack may arise that does not require physical harm, even though physical harm is
required for armed attacks caused by other sorts of weapons.” With the experience of the Dyn case, that time may be now.

The U.S. interpretation of the law would certainly seem to be open to such a finding. In the first place, the 2015 U.S. Department of Defense’s (DoD) *Law of War Manual* confirms in Chapter XVI (which addresses cyber operations) that the law of war applies to cyber, but admits that “[p]recisely how the law of war applies to cyber operations is not well-settled, and aspects of the law in this area are likely to continue to develop, especially as new cyber capabilities are developed and States determine their views in response to such development.”

Next, the *Law of War Manual* goes on to essentially incorporate the Koh approach by saying, “if cyber operations cause effects that, if caused by traditional physical means, would be regarded as a use of force under *jus ad bellum*, then such cyber operations would likely also be regarded as a use of force.” This intriguingly suggests that a use of force sufficient for *jus ad bellum* might exist even in the absence of physical injuries or destruction.

How? In listing examples of acts that could meet the use-of-force standard, the Manual says: “[c]yber operations [that] cripple a military’s logistics systems, and thus its ability to conduct and sustain military operations, might also be considered a use of force under *jus ad bellum*.” The footnote supporting this proposition points to a 1999 *Assessment of International Legal Issues in Information Operations* published by the DoD Office of the General Counsel. That document says:

> Even if the systems attacked were unclassified military logistics systems, an attack on such systems might seriously threaten a nation’s security. For example, corrupting the data in a nation’s computerized systems for managing its military fuel, spare parts, transportation, troop mobilization, or medical supplies may seriously interfere with its ability to conduct military operations. *In short, the consequences are likely to be more important than the means used.*

This illustrates that at least from DoD’s perspective, if a cyber event has significant enough consequences, it can be a *casus belli* even in the absence of physical injuries or destruction. The relevant question then would be: doesn’t an assault that caused “a stunning breach of global internet stability” and shut down half the internet qualify?

Complicating the issue is the July 2016 testimony before Congress by the State Department’s Coordinator for Cyber Issues Christopher Painter as to what he called “digital acts of war.” According to Painter, in determining on a “case-by-case, fact-specific” basis whether a cyber activity constitutes an “armed attack” sufficient to trigger the right of self-defense, “the actual or anticipated effects of a particular incident” are of “primary importance.” Painter says “the U.S. government believes that states should consider the nature and extent of injury or death to persons and the destruction of, or damage to, property.” If the cyber act “proximately” causes “death, injury, or significant destruction” it would “likely would be viewed as an armed attack.”

The problem, of course, is that while Painter’s formulation includes the obvious “death, injury, or significant destruction” standard, it doesn’t necessarily preclude finding that non-destructive cyber events could also produce “actual or anticipated effects” sufficient to permit an Article 51 response. It seems that Painter intentionally meant to be rather enigmatic as he also claims:

> As a general matter, states have not sought to define precisely (or state conclusively) what situations would constitute armed attacks in other domains, and there is no reason cyberspace should be different. In fact, there is a good reason not to articulate a bright line, as strategic ambiguity could very well deter most states from getting close to it.
Does calling a severe disruption “cyber vandalism” deter or incentivize?

While there may be a place for ambiguity in strategic deterrence, the Dyn case shows that it is not working for the U.S. The reason could well be the trivializing public characterization the government has been giving to events like the Dyn incident, not to mention applying similar language even when physical damage actually resulted. In the law, words do matter. Portraying something as “cyber vandalism” would not permit the U.S. to legally respond in the same way it could if it had been struck by a “physical bomb,” and that could have serious consequences for the development of deterrence in relation to cyber events.

Put another way, vandalism is ordinarily understood as a minor criminal law matter involving judicial processes, and not something that would sanction the use of force. As an international law matter, retorsion and countermeasures might be available – as may be other remedies under the law of state responsibility – but none of these options allow the use of force. However, activities that equate to “physical bombs” could readily be viewed as a national security threat where the response in the first instance could be a necessary and proportional use of force to counter them. To reiterate, the law enforcement paradigm suggested by “vandalism” is very different from the jus ad bellum/law of war regime that arises from national security threats, with the “law enforcement” response being much more limited.

Yet even where the cyber incident unquestionably fulfills the “physical damage” criteria, the U.S. inexplicably softens its classification. For example, in 2014, President Obama similarly used the term “cybervandalism” in denying that North Korea’s cyber operation against Sony pictures constituted an “act of war.” However, the Department of Defense Cyber Strategy document released in April of 2015 described the Sony incident much more seriously. It said (p. 2):

North Korea conducted a cyberattack against Sony Pictures Entertainment, rendering thousands of Sony computers inoperable and breaching Sony’s confidential business information. In addition to the destructive nature of the attacks, North Korea stole digital copies of a number of unreleased movies, as well as thousands of documents containing sensitive data regarding celebrities, Sony employees, and Sony’s business operations. North Korea accompanied their cyberattacks with coercion, intimidation, and the threat of terrorism. The North Korean attack on Sony was one of the most destructive cyberattacks on a U.S. entity to date. (Emphasis added.)

Likewise, Lisa Monaco, serving as Assistant to the President for Homeland Security and Counterterrorism, said in July of 2016 that the Sony attack “had crossed a threshold,” adding that it “was both destructive, it fried the computers of Sony Pictures, took them offline and it was coercive.” Given those facts as to physical destruction and coercion, it is hard to argue that the Sony attack did not meet the U.S. and, indeed, the world’s definition of “armed attack.”

Regarding last summer’s hack of thousands of Democratic National Committee (DNC) emails, Ms. Monaco emphasized the gravity of the event, calling it a “serious, serious issue, a serious thing if there is deliberate intrusion for the purpose of coercing and influencing the political process.” The distinctive nature of the target – the U.S. election system – caused John Brennan, Director of the Central Intelligence Agency, to conclude that “[o]bviously interference in the U.S. election process is a very, very serious matter.”

Despite the consensus about the seriousness and uniqueness of cyber efforts to interfere with the political process, the President again sought to downplay the incidents. In early September he acknowledged that the Russians have been attacking U.S. institutions on the internet” but has also said that:

Our goal is not to suddenly in the cyber arena duplicate a cycle of escalation that we saw when it comes to other arms races in the past, but rather to start instituting some norms so that everybody’s acting responsibly . . . . What we cannot do is have a situation in
which suddenly this becomes the wild, wild West, where countries that have significant cybercapacity start engaging in unhealthy competition or conflict through these means.

By early October the U.S. government was nevertheless explicitly accusing the Russian government of directing what it was calling “compromises” of (but not “attacks” on) cyber systems. It claimed that what it said were “thefts and disclosures” were “intended to interfere with the US election process.” Without referencing a legal basis, Josh Ernest, the White House Press Secretary, said a few weeks ago (but before the Dyn case) that there would be a “response” to these “thefts and disclosures.”

In doing so, Ernest added to the legal muddle by insisting that the response would be “proportional.” The response to a criminal matter like a “compromise” or “disclosure” or even a “theft” is a judicial one; a “proportional” response is, however, the language of force sounding in jus ad bellum, not law enforcement. Confusingly, he also said it “is unlikely that our response would be announced in advance” – again, jus ad bellum terms mixed with criminal law rhetoric. It is true, that an “armed attack” could also be a criminal offense, but the way it is being publicly presented suggests little cognizance of the critical differences between the two legal regimes, or the effect on deterrence those differences might have.

**Deterrence and dithering?**

Mr. Ernest further obfuscated the matter when he asserted that “[i]t’s certainly possible that the President could choose response options that we never announce.” This is hardly what would or should occur if it really is just vandalism – a criminal law matter – and it is not the way to go about deterring actors from similar behavior. How will people be deterred if the consequences are unknown? Adding to the confusion are press reports that suggest experts are not optimistic that the U.S. even has a clear vision of what the “proportional” response should be, announced or not.

Harvard Law Professor Jack Goldsmith, long a critic of what he calls the U.S.’s “feckless” cyber deterrence policy, warned that the U.S. government’s “dithering” in response to previous cyber incidents (including the 2015 Office of Personnel Management data breach that may have affected as many as 32 million people) was dangerous. He avowed that:

> Such a pattern of vacillation in response to very damaging cyber-operations will not deter our adversaries; it will embolden them. It will especially embolden them since the responses the USG finally settles on are much less than proportionate to the damage caused.

Susan Hennessy, a legal scholar at the Brookings Institute, differed somewhat with Goldsmith by asserting that U.S. deterrence policy has been successful to the extent that the U.S. “has never been the victim of a cyber attack that genuinely threatened lives.” She helpfully notes that the “Administration quietly released its policy on cyber deterrence late last year.” That policy states that “the Administration is most concerned about threats that could cause wide-scale disruption, destruction, loss of life, and significant economic consequences for the United States and its interests.” These, it says, would include (but are not limited to):

- Cyber attacks or other malicious cyber activity intended to cause casualties
- Cyber attacks or other malicious cyber activity intended to cause significant disruption to the normal functioning of U.S. society or government, including attacks against critical infrastructure that could damage systems used to provide key services to the public or the government.
• Cyber attacks or other malicious cyber activity that threatens the command and control of U.S. military forces, the freedom of maneuver of U.S. military forces, or the infrastructure on which the U.S. military relies to defend U.S. interests and commitments.

• Malicious cyber activity that undermines national economic security through cyber-enabled economic espionage or sabotage.

Hennessey believes that tampering with the mechanisms of the election is still a “below the threshold” activity (that is, below the “armed attack” standard), although she agrees that such actions combined with others might collectively exceed it. She also points to Phil Walters’ work describing a sophisticated Russian strategy to employ various “below established threshold activities” (BETA) and its relation to deterrence. She concludes that while the U.S.’s deterrence is working so as to be “effectively preventing very serious activity, at least for now,” its responses to BETA “are reactive and unpredictable, which undercuts the deterrent effect.” She closes by saying:

U.S. deterrence policy currently has the feeling of roulette. Maybe the house still wins overall, but it is clear that actors like Russia are happy to keep spinning the wheel while they’re ahead.

Indeed. Less than two weeks after Hennessey wrote her piece, an undeterred actor launched the gigantic Dyn assault that hobbled half of the Web. Even the Department of Homeland Security admitted just a month before the Dyn attack that the U.S. “has experienced increasingly severe and significant cyber incidents affecting both the private sector and Federal Government.” That admission, along with the new Dyn case, ought to make it clear that the U.S. needs to retool its cyber deterrence strategy.

**What to do? Clarifying terms**

Deterrence is a devilishly complex endeavor, especially where cyber is concerned – but clarifying the law can help. Shortly after the Dyn incident, Mr. Clapper lamented:

[W]e don’t have enough body of law yet. We haven’t, in my opinion — this is not company policy; it’s just me speaking — but we have not been able to generate either the substance or the psychology of deterrence in the cyber realm. And that’s going to continue to be an issue for us.

Irrespective as to whether Clapper is really correct about whether an adequate body of law exists to support deterrence, the fact is that there seems to be that perception among many U.S. officials. In truth, the law itself may not be the problem as much as the proper application of the law (and especially the U.S.’s view of it) to the facts. That proper application can be facilitated by cleaning up the language officials use about cyber incidents, and to synchronize it with announced U.S. interpretations.

To effectively deter, consistency and accuracy of language is indispensable. Since the U.S. has elected to characterize any use of force as sufficient to trigger a right to self-defense under Article 51, when events occur that plainly meet that standard (and even event that cross the more demanding “armed attack” threshold), then they need to be declared a use of force. For example, if the descriptions by DoD and government officials about the scope and intensity of the physical damage inflicted by North Korea in the Sony cyber incident are accurate, it quite obviously meets the standard established in the Koh speech, the **DoD Law of War Manual**, and Painter’s testimony.

Watering down official characterizations of the Sony attack (where computers were “fried” and “thousands” of them rendered “inoperable”) to merely being an incident of cybervandalism carries real consequences. At best confusion arises, and at worst, a norm develops that says to potential cyber adversaries that even if they inflict damage and coercion of that level of scale and intensity, they are still not in peril of anything worse than an indictment in a U.S. court that they can understandably expect will never result in an actual prosecution.
To be sure, there are acts that may appropriately be characterized as just cybervandalism. For example, in early 2015 when Islamic State hackers penetrated U.S. Central Command social media accounts, the U.S. branded it as “purely a case of cybervandalism.” Even though the hackers posted “threatening messages and propaganda videos, along with some military documents,” the command maintained that the “operation military networks were not compromised and there was no operational impact to U.S. Central Command.” It is possible that as cyber enables the ability to hyper-personalize threats on an unparalleled scale, the impact on civilians of hacking of social media accounts may eventually cause the development of an international norm prohibiting doing so. Still, every hostile cyber activity cannot and should not be characterized as a use of force, even under the U.S.’s more permissive standard.

It does help when, as noted above, the U.S. defines the cyber activities it wants to explicitly deter (beyond generically wanting to deter all crime). The problem with the listing is that it may include activities – cyber espionage for example – that are rightly violative of domestic U.S. law, but would not necessarily be something that the U.S. and its allies would want to be considered in international law, at least at the moment, as a casus belli. It may be suitable for the development of new norms not involving force in light of the enormous capability of cyber methodologies, but a clear delineation between what authorizes a forceful response, and what is limited to other options is what is needed.

In short, for deterrence to work there needs to be more precision in the official language used to describe specific incidents that comports to the U.S.’s own interpretation of a use of force that would authorize a response in self-defense. If the facts show an incident being characterized as a use of force sufficient to permit the use of force under Article 51, then the official language needs to be consistent with that assessment.

What to do? Develop norms as to the “red lines”

It is vitally important, however, to appreciate that simply because a particular cyber act may legally constitute an “armed attack” that might qualify for the political characterization of an “act of war,” that doesn’t mean that a country is obliged to respond to it with force. Indeed, there are many political reasons that would counsel against doing so. This is where Mr. Painter goes wrong with his discussion about “strategic ambiguity.”

In deterrence, ambiguity may be useful with respect to a response, but it is markedly less so when you are talking about the threshold. Misunderstandings as to where the proverbial “red lines” are set can lead to dangerous miscalculation, unintended escalation, and unwanted conflict. Given the enormous potential of cyber acts to do harm, potential actors ought not to get mixed messages as to how the U.S. considers harmful cyber activities.

Frustrations with the opaqueness as what cyber activity would constitute a casus belli appears to have motivated Congressman Mike Rounds to propose a bill earlier this year that would require the President to develop a policy for determining “when an action carried out in cyberspace constitutes an act of war against the U.S.”

Rounds points to testimony of Marine Lt. Gen. Vincent Stewart, director of the Defense Intelligence Agency, as part of his rational for the legislation. Stewart admitted that a “much fuller definition of the range of things that occur in cyber space [is needed], and then [we should] start thinking about the threshold where an attack is catastrophic enough or destructive enough that we define it as an act of war, I think that would be extremely helpful.”

Stewart isn’t alone in not “fully” understanding where the threshold lies. Other Pentagon leaders apparently are equally uncertain, something that raises the obvious question: if our leaders don’t know, how can we expect potential adversaries to understand which acts might spark a full-blown war? At the same time, except in the most aggravated cases, enumerating in advance precisely which cyber acts exceed the use of force threshold might be nearly impossible.
This is where norm development in international law comes into play. In doing so, the U.S. needs to use the language of international law. Political terms like “digital acts of war” are unhelpful not only because they do not track with the language of the law, they also can imply to the general public a level of response that is unnecessarily provocative and even inconsistent with the proportionality and necessity factors intrinsic to a lawful exercise of self-defense, especially in the complex cyber arena.

As discussed, the U.S. has, in fact, laid out in broad terms what kind of cyber activities it wants to deter, and generally how it interprets the law applicable to cyber operations. What is required now is for the U.S. to act consistently with these conceptual positions when cyber incidents actually occur. We now seem to be in a cycle where we are facing ever more dangerous and damaging cyber incidents, yet they are rarely given the appellations established U.S. legal interpretations would seem to indicate. Instead, incidents too often are characterized with language that would put them outside the kinds of activities that would authorize a forceful Article 51 response.

The U.S. also has to be more forthright about its response to incidents because that too influences norm development. True, there may be times, as the White House spokesman Josh Ernest said, that the U.S. would “never announce” a response to a particular cyber incident, but that should very much be the exception and not the rule. As Bloomberg News’ Eli Lake argued last July after the DNC hack:

[T]here is also a consequence for keeping quiet. It might give Russian hackers the impression that the U.S. is uninterested in deterring them. Indeed, it appears they are under that impression already.

Transparency should not be underestimated as a deterrence factor. Potential cyber attackers calculate exactly what kind of malicious activity will generate a response, and how costly that response might be.

The bigger picture

It is crucial that the U.S. express its positions unmistakably about cyber incidents it has suffered, particularly given the approach of two of the world’s most formidable cyber actors. Professor Schmitt noted in 2014 that:

The UN Group of Governmental Experts, which includes representatives from Russia and China, agreed in 2013 that international law applies to cyberspace. Interestingly, Russia and China did not agree to a reference to international humanitarian law and China reportedly does not accept the applicability of IHL in cyberspace. (Emphasis added and citations omitted)

For example, the Chinese acknowledge that “although the existing laws on armed conflicts and general international principles may all apply to cyberspace, there are still many issues that need clarification...[t]he international community should, therefore, revise existing laws – but it is important that this international legal framework maintains sufficient openness and flexibility.”

Although purportedly not officially speaking for the Chinese government, Professor Huang ZhiXiong of China’s Wuhan University Institute of International Law is reported to have said at a 2015 conference on cyber law that:

In his view, the Tallinn factors relevant to evaluating when a cyber activity rises to a use of force (which include severity, directness, and invasiveness) are too malleable and the bar for what activities are uses of force should be higher. Second, he sought a higher bar than Tallinn 1.0 sets for when a state may invoke the right of self-defense. In his view, a state does not have a right of self-defense against attacks by non-state actors, nor does a state have the right of self-defense against an imminent attack.

If the Chinese government de facto adopts (or has already adopted) Professor ZhiXiong’s perspective as to the inapplicability of the right to self-defense in cyber incidents, and Russia fuses with that view,
their combined impact would be very influential in the development of an international norm that is contrary to the U.S. view.

**Some final thoughts**

It is inarguable that the U.S. needs to be judicious in its characterizations of, and response to, cyber events. No one wants to unnecessarily aggravate an already difficult situation. Uncertainty as to how to effectively respond and still avoid counterproductive escalation are real problems of deterrence. But before determining whether and how to respond, the legal options need to be apparent. In that regard the U.S. is at the point where it needs to be more forthright when incidents occur that appear to violate its own announced standards as to when a cyber action equates to an “armed attack.”

Again, calling something the equivalent of an “armed attack” so as to permit a forceful and proportional response in self-defense under Article 51 does not mean that such action necessarily be forthcoming in every instance. Rather, what it would do is make it unmistakable to all concerned that the U.S. asserts it has a lawful option to use force in self-defense if it chooses to do so, not that it will in each case.

When the U.S. fails to properly characterize cyber incidents, and frequently suggests that they are simply vandalism, thefts or other matters which are readily interpreted by cyber actors and publics around the world as being within the law enforcement modality and outside of the *jus ad bellum* legal regime, no one should be surprised if norms begin to emerge more in keeping with what Russia, China, and hostile cyber actors prefer.

Deterrence in the cyber realm quite obviously needs strengthening, and dealing with the legal piece of that effort matters. We still have the chance to set the record straight – to develop that “body of law” Director Clapper believes we are missing – but that opportunity diminishes with each passing incident where the proper legal characterization is understated and muddled.
Respect for Sovereignty in Cyberspace

Michael N. Schmitt* and Liis Vihul**

I. Discord Regarding Sovereignty

In the late 1990s, the international legal community’s attention began to turn to a new form of warfare, then labeled “computer network attack,” a type of information operations.1 At the time, the Department of Defense (DoD) was at the cutting edge of thought regarding the legal significance of these operations. By 1999 its consideration of the issue had matured, and the Office of the General Counsel released An Assessment of International Legal Issues in Information Operations,2 which considered the application of the *jus ad bellum* and *jus in bello* rules, space law, international telecommunication law, the law governing espionage activities, specific treaty regimes, and domestic law to military operations in cyberspace. *Information Operations* operated from the premise that international law applies in cyberspace. This remains the U.S. approach nearly two decades later.3

Yet, the document was cautionary. As it perceptively noted, the international legal system is reactive in the sense that it typically develops in

---

*I* Professor of International Law, University of Exeter; Chairman, Stockton Center for the Study of International Law, U.S. Naval War College; Francis Lieber Distinguished Scholar, Lieber Institute, U.S. Military Academy at West Point; Director, Tallinn Manual 2.0 Project. The views expressed are those of the author in his personal capacity.

**CEO, Cyber Law International; Managing Editor, Tallinn Manual 2.0 Project.

1. See Joint Chiefs of Staff, Joint Publ’n 3-13, Joint Doctrine for Information Operations viii, GL-5 (1998), http://www.c4i.org/jp3_13.pdf [https://perma.cc/F6LP-T4UJ] (approving the addition of “computer network attack” to the Department of Defense Dictionary of Military and Associated Terms). Information operations are “[t]he integrated employment, during military operations, of information-related capabilities in concert with other lines of operation to influence, disrupt, corrupt, or usurp the decision-making of adversaries and potential adversaries while protecting our own.” Joint Chiefs of Staff, Joint Publ’n 1-02, Department of Defense Dictionary of Military and Associated Terms 110 (2016), https://fas.org/irp/doddir/dod/jp1_02.pdf [https://perma.cc/7WWV-NHYK].


response to particular situations and their consequences.\textsuperscript{4} This being so, the assessment warned, “we can make some educated guesses as to how the international legal system will respond to information operations, but the direction that response actually ends up taking may depend a great deal on the nature of the events that draw the nations’ attention to the issue.”\textsuperscript{5} Evolution in the law’s interpretation in the cyber context was therefore inevitable.

What appears to have changed since then is the DoD’s position on sovereignty in cyberspace. In 1999, the question was not whether a State could violate another State’s sovereignty as a matter of law; rather, the challenge was identifying when cyber operations do so. That the prohibition on violation of sovereignty is a substantive rule of international law was an assumption that permeated the assessment. For example, it noted that in air law the entry by one State’s aircraft into another’s national airspace was “regarded as a violation of its sovereignty and territorial integrity.”\textsuperscript{6} In the maritime environment, the document pointed to the 1949 \textit{Corfu Channel} case,\textsuperscript{7} in which the International Court of Justice held that the penetration of Albanian territorial waters by British warships, and the minesweeping operation therein, without legal justification amounted to a violation of Albanian sovereignty.\textsuperscript{8}

Regarding cyber operations, the document observed that “[a]n unauthorized electronic intrusion into another nation’s computer systems may very well end up being regarded as a violation of the victim’s sovereignty. It may even be regarded as equivalent to a physical trespass into a nation’s territory . . . .”\textsuperscript{9} And with respect to responding by cyber means against individuals or groups operating from other States, it noted that:

\textit{[e]ven if it were possible to conduct a precise computer network attack on the equipment used by such individual actors, the state in which the effects of such an attack were felt, if it became aware of it, could well take the position that its sovereignty and territorial integrity had been violated.\textsuperscript{10}}

Thus, as framed in the 1999 DoD assessment, certain State cyber operations against other States might violate the latter’s sovereignty, that is, constitute an “internationally wrongful act.”\textsuperscript{11} In the same vein, and over a

\textsuperscript{4} Information Operations, supra note 2, at 464.
\textsuperscript{5} \textit{id}. at 465.
\textsuperscript{6} \textit{id}. at 464.
\textsuperscript{7} \textit{id}. at 481.
\textsuperscript{9} Information Operations, supra note 2, at 485.
\textsuperscript{10} \textit{id}. at 488.
\textsuperscript{11} An internationally wrongful act of a State consists of an action or omission that is attributable to the State under international law and constitutes a breach of an international
decade later, the premise of sovereignty as a primary rule of international law capable of being violated was accepted unanimously by the international law scholars and practitioners who prepared the 2013 *Tallinn Manual on the International Law Applicable to Cyber Warfare*, as well as those who produced its 2017 successor, *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*.\(^{12}\)

Recently, the DoD has indicated that it may have reassessed its position that sovereignty can be violated as a matter of international law in the cyber context. The prospect surfaced publicly in a panel presentation by Colonel Gary Corn, the Staff Judge Advocate of U.S. Cyber Command, at the 2016 “CyCon U.S.” conference.\(^{13}\) Then, on the day before the President’s inauguration, Jennifer O’Connor, the Department’s General Counsel, issued a memorandum titled “International Law Framework for Employing Cyber Capabilities in Military Operations” that dealt with, inter alia, the subject of sovereignty.\(^{14}\)

Addressed to the Commanders of the Combatant Commands and very senior lawyers throughout the DoD, the memorandum was initially unclassified and circulated widely internationally. However, it was later designated as “for internal use only,” and distribution is now restricted.\(^{15}\) Nevertheless, Corn and former Principal Deputy General Counsel of the DoD Robert Taylor have since published on the subject.\(^{16}\) Considering their positions as, respectively, the most senior legal advisor for the U.S. organization that engages in military cyber operations, the author of the memorandum, and a highly placed DoD attorney at the time it was issued, it

---


12. **TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS** 1 (Michael M. Schmitt & Liis Vihul eds., 2017) [hereinafter **TALLINN MANUAL 2.0**]; **TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE** (Michael N. Schmitt ed., 2013) [hereinafter **TALLINN MANUAL 1.0**]. Primary rules are those which impose either obligations or prohibitions on States. They must be distinguished from secondary rules of international law, that is, “the general conditions under international law for the State to be considered responsible for wrongful actions or omissions, and the legal consequences which flow therefrom.” Responsibility of States for Internationally Wrongful Acts, [2001] 2 Y.B. Int’l L. Comm’n 31, U.N. Doc. A/CN.4/SER.A/2001/Add.1. Examples of secondary rules include those regarding attribution and the remedies that are available to States when international law obligations owed them are breached.

13. Colonel Corn was, however, speaking in his personal capacity. The authors spoke on the same panel.


15. As one of the authors is a DoD employee, the document cannot be quoted in this article.

is reasonable to assume that their views are consistent with the DoD’s position.

By their approach, sovereignty does not operate as a rule of international law, the violation of which results in international legal responsibility.\textsuperscript{17} Instead, it is a “baseline principle . . . undergirding binding norms,”\textsuperscript{18} particularly the U.N. Charter Article 2(4) prohibition on the use of force and the customary international law prohibition on coercive intervention.\textsuperscript{19} This article examines the point of contention between the DoD’s earlier view, as well as the Tallinn Manuals’, and that which now appears to be the revised DoD position. Part II assesses the legal logic underlying the argument against the existence of such a rule and sets forth the position of the authors on the matter. Drawing on the approach adopted in Tallinn Manual 2.0, it focuses on territorial sovereignty and its inviolability by other States. In Part III, evidence that the prohibition on violating sovereignty reflects customary international law is surveyed. Included are discussions of treatment of the matter by international tribunals, States, and international organizations. A brief illustration of how the two views might play out in practice is offered in Part IV, together with the authors’ thoughts on the possible consequences of the debate.

II. Assessing the Argument Against a Primary Rule on Violations of Sovereignty

As noted, the authors of the two Tallinn Manuals unanimously agreed that the principle of sovereignty prescribes certain cyber operations conducted by States against other States. Tallinn Manual 2.0 accordingly provides in Rule 4 that “[a] State must not conduct cyber operations that violate the sovereignty of another State.”\textsuperscript{20} Corn took issue with the substance of the rule in a Just Security post that followed publication of the Manual and was subsequently joined by Taylor in an AJIL Unbound piece further developing the position.\textsuperscript{21}

Much of the argument they put forth is uncontroversial. For instance, both sides of the debate agree that the principle of sovereignty is the basis for the international law prohibitions of intervention and use of force.\textsuperscript{22} Yet, advocates of the “sovereignty-as-principle-only” approach draw the line at these two internationally wrongful acts, rejecting any directly operative

\begin{flushright}
\textsuperscript{17} Corn & Taylor, \textit{supra} note 16.
\textsuperscript{18} Corn, \textit{supra} note 16.
\textsuperscript{19} U.N. Charter art. 2, ¶4 (“All Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations.”).
\textsuperscript{20} \textit{TALLINN MANUAL} 2.0, \textit{supra} note 12, at 17 (Rule 4).
\textsuperscript{21} Corn, \textit{supra} note 16; Corn & Taylor, \textit{supra} note 16.
\textsuperscript{22} Corn & Taylor, \textit{supra} note 16; \textit{TALLINN MANUAL} 2.0, \textit{supra} note 12, at 11–12.
\end{flushright}
effect of the principle itself, such as a rule prohibiting the breach of territorial inviolability. According to Corn and Taylor:

[I]t is widely recognized that states have unquestioned authority to prohibit espionage within their territory under their domestic laws, but it is also widely recognized that international law does not prohibit espionage. States have long engaged in espionage operations that involve undisclosed entry and activities within the territory of other states, subject only to the risk of diplomatic consequences or the exercise of domestic jurisdiction over intelligence operatives if discovered and caught. Within this framework, it is understood that espionage may violate international law only when the modalities employed otherwise constitute a violation of a specific provision of international law, such as an unlawful intervention or a prohibited use of force. Thus states conduct intelligence activities in and through cyberspace, and generally, “to the extent that cyber operations resemble traditional intelligence and counter-intelligence activities . . . such cyber operations would likely be treated similarly under international law.” This framework applies equally to cyber operations directed at terrorist cyber infrastructure located within the territory of another state.

Further, the differences in how sovereignty is reflected in international law with respect to the domains of space, air, and the seas further supports the view that sovereignty is a principle, subject to adjustment depending on the domain and the practical imperatives of states rather than a hard and fast rule. For instance, in the case of the space domain, objects in orbit are beyond the territorial claims of any nation, and outer space – including outer space above another state’s territory – is available for exploitation by all. In the case of the air domain, the regime is highly restrictive, such that any unconsented entry into the airspace of another state is regarded as a serious violation of international law subject to such exceptions as self-defense, Security Council authorization, or force majeure. In the case of the seas, many entries into and travels through the territory of another state are permissible without the consent of that state, but there are conditions under which such entry would be a violation of international law – it depends on the particular facts and circumstances. The fact that states have developed vastly different regimes to govern the air, space, and maritime domains underscores the fallacy of a universal rule of sovereignty with a clear application to the domain of cyberspace. The principle of sovereignty is universal, but its application to the unique particularities of the cyberspace domain remains for states to determine through state practice and/or the development of treaty rules.23

23. Corn & Taylor, supra note 16 (emphasis omitted).
This is where their argument breaks down, for it fails to recognize that each of the legal regimes cited—air, space, maritime, and that governing espionage—are premised on territorial integrity and inviolability. Regarding the air domain, consider a Russian military aircraft that briefly “cuts the corner” into Estonian airspace. There is no State practice supporting treatment of these incidents, which are the subject of the ongoing NATO Baltic Air Policing mission, as a use of force or coercive intervention.\(^\text{24}\) On the contrary, they constitute violations of Estonian national airspace,\(^\text{25}\) and thereby its territorial sovereignty. As will be seen, this is the generally consistent approach States take to aerial intrusions into inviolable national airspace.\(^\text{26}\)

With respect to outer space, States have confirmed in treaty law that it is not subject to national appropriation by claim of sovereignty.\(^\text{27}\) This indicates that but for that rule, which is now accepted as customary in nature, territorial sovereignty would by default be viewed as extending beyond the airspace above a State’s sovereign territory into outer space. Space law is therefore *lex specialis* that allows, for instance, States to place space objects into geostationary orbit above the subjacent territory of other States.\(^\text{28}\) It is a legal accommodation agreed to by States that is designed to permit them to operate in outer space in ways that might otherwise be prohibited through application of the *lex generalis* rules of territorial sovereignty.

The law of the sea also supports the primary-rule status of territorial sovereignty. Recall that Corn and Taylor opine, in reference to maritime activities, that “many entries into and travels through the territory of another


\(^{25}\) See, e.g., Estonia Says Russian Aircraft Violated Airspace Again, RADIO FREE EUROPE/RADIO LIBERTY (Sept. 6, 2016), http://www.rferl.org/a/russia-estonia-airspace-violated/27970888.html\[https://perma.cc/9JCW-3A3Q\] (recounting the Estonian military’s claim that a Russian aircraft violated Estonian airspace by flying within it “without permission for about 90 seconds”).

\(^{26}\) The Chicago Convention provides, “The contracting States recognize that every State has complete and exclusive sovereignty over the airspace above its territory.” Convention on International Civil Aviation, art. 1, Dec. 7, 1944, 61 Stat. 1180, 15 U.N.T.S. 295. Use of the term “recognize” confirms the customary international law character of such sovereignty.


\(^{28}\) See, e.g., Definition and Delimitation of Outer Space and the Character and Utilization of the Geostationary Orbit, 2001 DIGEST OF UNITED STATES PRACTICE IN INTERNATIONAL LAW, ch. 12, § C(4) at 722 (“Article II . . . further states that outer space is not subject to national appropriation by claim of sovereignty or by any other means. Thus, a signatory . . . cannot appropriate a position in the [geostationary orbit] either by claim of sovereignty or by means of use, or even repeated use, of such an orbital position.”).
State are permissible without the consent of that State, but there are conditions under which such entry would be a violation of international law—it depends on the facts and circumstances. While their statement of the law is correct, the authors fail to acknowledge the reason why consent of the coastal States need not be obtained when another State’s vessel wishes to sail through the former’s territorial sea. States have long enjoyed territorial inviolability vis-à-vis their coastal waters. The regimes of innocent, transit, and archipelagic passage developed as customary and treaty-law exceptions to the territorial sea’s inviolability; they modify the baseline principle that maritime borders may not be pierced by other States. Territorial inviolability remains intact, subject to the exceptions.

Finally, the issue of espionage can also be viewed through the prism of territorial sovereignty. Corn and Taylor point to the long-standing State practice of engaging in espionage activities on foreign territory, which they suggest is not viewed by States as a violation of international law. Although they do not set forth the legal basis for this conclusion, a plausible argument supporting it is that, based on the extensive State practice of conducting espionage abroad, espionage constitutes a customary exception to the general rule that territorial sovereignty is inviolable. The weakness in this rationale is the limited amount of opinio juris on point, for a new customary international law rule must be grounded in both State practice and opinio juris.

By the opposing view, espionage on another State’s territory is de jure a violation of that State’s territorial sovereignty. For those advocating this position, the question in the case of remotely conducted cyber-espionage operations, therefore, would be identical to that which must be asked of any other cyber operation—at what point does an operation that does not entail physical presence on another State’s territory qualify as a violation of territorial sovereignty? The manner in which the Tallinn Manual 2.0 answers this question is set out below. But irrespective of which side of the debate

29. Corn & Taylor, supra note 16.
31. This conclusion is without prejudice to authorization or mandate by the U.N. Security Council under Chapter VII of the U.N. Charter, operations conducted pursuant to the right of self-defense, or situations provided for in the law of the sea, such as force majeure or distress. See U.N. Charter arts. 42 (providing the basis for “peace enforcement” operations), 51 (affirming the right of self-defense); Law of the Sea Convention, supra note 30, art. 18, ¶ 2 (allowing stopping and anchoring in territorial seas as rendered necessary by exigent circumstances).
32. Corn & Taylor, supra note 16.
33. See, e.g., TALLINN MANUAL 2.0, supra note 12, at 18–19, 171 (noting that “[i]n the cyber context . . . it is a violation of territorial sovereignty for an organ of a State, or others whose conduct may be attributed to the State, to conduct cyber operations while physically present on another State’s territory,” and suggesting the majority view is that cyber espionage would constitute violation of sovereignty if the individual committing the espionage operation “is on another State’s territory while nonconsensually engaging in the operation”).
one takes, territorial sovereignty resides at the heart of the underlying legal logic.

To bolster their position on territorial sovereignty, Corn and Taylor turn to the work of scholars, principally Ian Brownlie’s classic work, *International Law.* It is true that Brownlie characterizes the term “sovereignty” as “rather descriptive in character, referring in a ‘catch-all’ sense to the collection of rights held by a state.” What they fail to note, however, is that Brownlie, citing *Corfu Channel*, undeniably sees territorial inviolability as one of those rights and observes that other States accordingly shoulder a “correlative duty of respect for territorial sovereignty.”

The seminal treatise in the field, Lassa Oppenheim’s *International Law*, also endorses the notion that territorial sovereignty must be respected and that failure to do so constitutes a violation of international law. This view was advanced in the book’s first edition, published in 1905.

The duty of every State to abstain itself and to prevent its organs and subjects from any act which contains a violation of another State’s independence or territorial and personal supremacy is correlative to the respective right of the other State. It is impossible to enumerate all such actions as might contain a violation of this duty. But it is of value to give some illustrative examples. . . . Further, in the interest of the territorial supremacy of other States, a State is not allowed to send its troops, its men-of-war, and its police forces into or through foreign territory, or to exercise an act of administration or jurisdiction on foreign territory, without permission.

It has stood the test of time, for, although revised to accommodate new factual circumstances and the maturation of international law, the analysis was maintained by each of the book’s distinguished subsequent editors.

The most recent edition (9th), published in 1992, provides,

---

34. Corn & Taylor, *supra* note 16.
35. JAMES CRAWFORD, BROWNlie’S PRINCIPLES OF PUBLIC INTERNATIONAL LAW 448 (8th ed. 2012).
36. Id.; *see also* James Crawford, Sovereignty as a Legal Value, in *INTERNATIONAL LAW* 117, 121 (James Crawford & Martti Koskenniemi eds., 2012) (“As a general matter, [sovereign] authority is exclusive: normally, governmental activity carried out on the territory of another state is only lawful if performed there with the latter’s consent . . . .’’); H.W. Halleck, *International Law* 270 (1861) (“Every right has its correlative duty,’ which in the present context would mean that a State’s right to exclusive authority within its territory carries with it the correlative duty to respect the same right of other States); Malcolm N. Shaw, *International Law* 353 (7th ed. 2014) (“The principle of respect for the territorial integrity of states is well founded as one of the linchpins of the international system, as is the norm prohibiting interference in the internal affairs of other states.”).
All states are under an international legal obligation not to commit any violation of the independence, or territorial or personal authority, of any other state.

It is not feasible to enumerate all such actions as might constitute a breach of a state’s duty not to violate another state’s independence or territorial or personal authority. But it is useful to give some illustrative examples. A state is not allowed to send its troops, its warships, or its police forces into or through foreign territory, or its aircraft over it, or to carry out official investigations on foreign territory or to let its agents conduct clandestine operations there, or to exercise an act of administration or jurisdiction on foreign territory, without permission.

As is apparent, it is misguided to assert that there must exist a cyber-specific rule for cyber operations not amounting to a wrongful use of force or coercive intervention, but manifesting on another State’s territory, to qualify as violations of territorial sovereignty. The pressing task is, instead, to identify the criteria for violation thereof by means of cyber operations. Only if *lex specialis* subsequently emerges through treaty or crystallization of customary law, as in the case of outer space, will cyber operations that would otherwise violate a State’s territorial sovereignty be permissible.

Treating violations of sovereignty as a primary rule of international law, *Tallinn Manual 2.0* seeks to add granularity to the circumstances in which a cyber operation might violate a State’s territorial sovereignty. The commentary to Rule 4, set out above, provides that, as a general matter, “[c]yber operations that prevent or disregard another State’s exercise of its sovereign prerogatives constitute a violation of such sovereignty and are prohibited by international law.”

States enjoy sovereignty over cyber infrastructure, persons, and cyber activities located on their territory. This includes both public and private cyber infrastructure. For the experts who produced *Tallinn Manual 2.0*, the difficulty lay in identifying those cyber operations that would violate it. They conducted their analysis along two axes: “(1) the degree of infringement upon the target

40. TALLINN MANUAL 2.0, supra note 12, at 17.
41. Id. at 13 (Rule 2).
42. Id. at 13–14. This is without prejudice to exceptions provided for in law, such as diplomatic protection. See, e.g., id. at 209 (“Premises of a mission’ refers to ‘the buildings or parts of buildings and the land ancillary thereto, irrespective of ownership, used for the purposes of the mission.”) (quoting Vienna Convention on Diplomatic Relations, art. 1(i), Apr. 18, 1961, 23 U.S.T. 3227, 500 U.N.T.S. 95); id. at 212 (“Cyber infrastructure on the premises of a diplomatic mission or consular post is protected by the inviolability of that mission or post.”).
State’s territorial integrity; and (2) whether there has been an interference with or usurpation of inherently governmental functions.\footnote{Id. at 20.}

With respect to infringement on territorial integrity, there was consensus that a State’s cyber operation causing physical damage or injury on the territory of another State violates the latter State’s territorial sovereignty. The group also concurred that a cyber operation resulting in a loss of functionality (such that the targeted cyber infrastructure or the equipment upon which it relies needs to be repaired or replaced) qualifies as a violation. No consensus could be achieved, however, as to remote cyber operations generating other consequences. Some experts treated the aforementioned consequences as the threshold for violation. Others suggested that violations of sovereignty might include additional operations but were unable to agree upon a definitive threshold to apply.\footnote{Id. at 20–21.}

Clearly, however, not all cyber operations that manifest, either partially or totally, on another State’s cyber infrastructure infringe that State’s territorial inviolability.\footnote{See Oppenheim’s International Law, supra note 39, at 385 ("However, not all acts performed by one state in the territory of another involve a violation of sovereignty.").} As an example, the transmission of propaganda by one State into other States from platforms in outer space or on the high seas is not considered to violate sovereignty, even when done against the target States’ wills.\footnote{Bruce A. Hurwitz, The Legality of Space Militarization 29–30 (1986).} The examples cited by Corn and Taylor would generally fall into this category. It is correct that cyber operations involving “cyber effects in, yet invisible to, the territorial State, but that only manifest operationally in the area of hostilities”\footnote{Corn & Taylor, supra note 16.} are generally permissible. Similarly, “[w]here the proposed cyber action is focused solely against the individual accounts or facilities of terrorists or terrorist organizations widely recognized as such, and when the cyber actions will generate only de minimis effects on nonterrorist infrastructure within the host State, international law does not preclude those cyber actions.”\footnote{Id. at 7.} Yet, citing select examples of cyber operations that States are unlikely to consider violations of territorial sovereignty does not disprove the existence of a primary rule prohibiting breaches of territorial inviolability in other cases. On the contrary, it demonstrates the need to develop interpretive criteria by which that rule will be applied.

Tallinn Manual 2.0 additionally notes that a violation of sovereignty occurs whenever a cyber operation interferes with or usurps another State’s inherently governmental functions.\footnote{Tallinn Manual 2.0, supra note 12, at 21.} The experts found it difficult to define “inherently governmental functions” with granularity. \textit{Id.} at 22. However, certain functions clearly
fact that, pursuant to the notions of internal and external sovereignty, these functions fall within the exclusive purview of the State. Such violations need not be accompanied by any damage or injury, and unlike the prohibition on intervention into a State’s *domaine réservé*, no coercive intent or effect is required. However, as the focus of the debate over sovereignty is on its territorial aspect, the discussion that follows shall be limited to territorial sovereignty and its inviolability.

The *Tallinn Manual 2.0* approach to sovereignty appears to be widely shared. Little criticism of the “sovereignty-as-rule” position, which was also reflected in the first edition of the *Tallinn Manual*, was heard during the nearly four years between publication of the two editions. On the contrary, discussion of sovereignty in the cyber context surrounded the identification of those cyber activities that might violate another State’s sovereignty.

Additionally, a draft of the *Tallinn Manual 2.0* rule on violation of sovereignty and its accompanying commentary was discussed in three meetings of over fifty States and international organizations that were convened by the Dutch Ministry of Foreign Affairs in 2015 and 2016. Many of the States subsequently provided voluminous unofficial written comments. They voiced no meaningful objection to Rule 4. Instead, the comments focused on application of the rule to specific situations. There was even consideration of whether the prohibition encompassed cyber activities by non-State groups, a view acknowledged, but not accepted, in the *Manual*. Throughout this process, it appeared to be received knowledge that a primary rule on territorial-sovereignty violations existed and applied to cyber operations.

### III. Evidence of a Primary Rule on Violations of Sovereignty

The question at hand is whether the principle of sovereignty operates as a primary rule of customary international law, imposing an obligation on States to respect the inviolability of other States’ territories. If so, it

---

50. *Id.* at 13 (Rule 2); *id.* at 16 (Rule 3).
imposes significant operational limits on State activities on, or with effects in, the territory of those States.

In the view of the authors, overwhelming evidence of State practice and opinio juris—the foundational elements of customary international law—supports the assertion that a primary rule not to violate the territorial sovereignty of other States exists. Examples of such practice and opinio juris are offered below. Additionally, pursuant to Article 38(1)(d) of the Statute of the International Court of Justice, “judicial decisions and the teachings of the most highly qualified publicists of the various nations” constitute “subsidiary means for the determination of rules of law.”

Since judicial decisions, in particular those of the International Court of Justice, are especially persuasive subsidiary means for assessing whether a customary law rule has crystallized, the examination of the supporting evidence begins with an appraisal of a number of key cases. As to the work of highly qualified publicists (scholars), the scholarship cited earlier self-evidently qualifies as such. Significant in the cyber context are the two Tallinn Manuals, the collective work of nearly forty scholars, many of whom are internationally renowned. They too would meet the requirements of Article 38(1)(d). Although length constraints preclude a comprehensive catalogue of support for the existence of a primary rule on sovereignty, that which is set forth below is proffered regarding the substance of the norm as well as to indicate the breadth and depth of the corroborating evidence.

A. Judicial Treatment

The premise that it is unlawful for a State to act on the territory of another State without the latter’s consent has long been recognized by international tribunals. In the 1927 *Lotus Case*, the Permanent Court of International Justice observed that “the first and foremost restriction imposed by international law upon a State is that—failing the existence of a permissive rule to the contrary—it may not exercise its power in any form in

(“It is of course axiomatic that the material of customary international law is to be looked for primarily in the actual practice and opinio juris of States . . . .”). The classic case addressing these requirements is North Sea Continental Shelf (Ger. v. Den.; Ger. v. Neth.), Judgment, 1969 I.C.J. Rep. 3 (Feb. 20). For further discussion, see INT’L LAW ASS’N, COMM. ON FORMATION OF CUSTOMARY (GEN.) INT’L LAW, STATEMENT OF PRINCIPLES APPLICABLE TO THE FORMATION OF GENERAL CUSTOMARY INTERNATIONAL LAW ¶ 10 (2000); Yoram Dinstein, *The Interaction Between Customary International Law and Treaties*, in RECUEIL DES COURS 322 (2006).


57. ICJ Statute, *supra* note 54, art. 38(1)(d); see TALLINN MANUAL 2.0, *supra* note 12, at xii–xxii (listing scholars from various nations contributing to both the *Tallinn Manual 2.0* and *Tallinn Manual 1.0*).
the territory of another State.”58 In other words, the court treated the principle as one that sets binding limits on a State’s activities on foreign territory; when a State acts without the territorial State’s consent, the former is in breach of an obligation owed the latter to respect its sovereignty.

This view of the law has been adopted by the Permanent Court’s successor, the International Court of Justice. Indeed, in its first case, Corfu Channel, the court dealt with accusations of violations of sovereignty.59 The case involved an incident in which British warships passing through the Corfu Channel in Albanian territorial waters in 1946 struck naval mines.60 Following the incident, the Royal Navy again sailed through the waters, this time to conduct minesweeping operations.61 The United Kingdom sought a finding that Albania was responsible for the damage to two of its vessels and the ensuing loss of life, and an order that it pay compensation.62 Albania counterclaimed, asking the court to decide whether the “United Kingdom under international law violated the sovereignty of the Albanian People’s Republic by reason of the acts of the Royal Navy in Albanian waters,” and, if so, whether there was a duty to provide satisfaction.63

The court held Albania responsible for the damage and loss of life on the basis that it had failed to warn the United Kingdom of the dangers posed by transit through the Corfu Channel.64 More important to the territorial-sovereignty issue were the findings of the court relative to Albania’s claim. The United Kingdom did not contest Albania’s sovereignty over the waters, nor did it suggest the absence of a norm precluding violations of sovereignty. Instead, it argued that a special maritime legal regime, innocent passage, allowed for transit through international straits lying in a State’s territorial sea, even in the absence of consent.65 The court agreed and therefore was “unable to accept the Albanian contention that the Government of the United Kingdom ha[d] violated Albanian sovereignty by sending the warships through the Strait without having obtained the previous authorization of the Albanian Government.”66 The waters were subject to territorial sovereignty, but an exception applied.

60. Id. at 12–13.
61. Id. at 13.
62. Id. at 10–11.
63. Id. at 6.
64. Id. at 23.
65. Id. at 27. As it is an international strait, under the modern law of the sea, passage through the Corfu Channel would be “transit passage.” Law of the Sea Convention, supra note 30, arts. 37–38.
An opposite conclusion was reached with respect to the minesweeping. Because the operations were conducted without Albania’s consent, and no exception operated, the court concluded that:

Between independent States, respect for territorial sovereignty is an essential foundation of international relations. The Court recognizes that the Albanian Government’s complete failure to carry out its duties after the explosions, and the dilatory nature of its diplomatic notes, are extenuating circumstances for the action of the United Kingdom Government. But to ensure respect for international law, of which it is the organ, the Court must declare that the action of the British Navy constituted a violation of Albanian sovereignty.67

Since the 1949 Corfu Channel judgment, the International Court of Justice has continued to address the issue of, and often find, internationally wrongful violations of sovereignty. In 1973, it considered the legality of French atmospheric nuclear testing in the South Pacific.68 The case, Nuclear Tests, involved an Australian request for a declaratory judgment that the French testing violated international law, as well as a permanent order prohibiting France from carrying out further tests.69 Although it was dismissed on procedural grounds, what is relevant to the issue of breach of sovereignty as a primary rule is the Australian government’s position that the “deposit of radio-active fall-out on the territory of Australia and its dispersion in Australia’s airspace without Australia’s consent . . . violates Australian sovereignty over its territory.”70

In its Memorial, Australia set forth its legal logic in making the claim:

The Government of Australia repeats that its case rests upon several bases: on the mere fact of trespass, on the harmful effects associated with trespass, and on the impairment of its independent right to determine what acts shall take place within its territory. In this connection, the Government of Australia wants to emphasize that the mere fact of trespass, the harmful effects which flow from such fall-out and the impairment of its independence, each clearly constitute a violation of the affected State’s sovereignty over and in respect of its territory.71

The court then addressed the issue of a legal right to allege a violation of sovereignty.

The evident character of Australia’s legal interest in a claim alleging violation of its sovereignty over and in respect of its territory is such

67. Id. at 35.
69. Id. ¶ 26.
as to make any extended argument upon this point superfluous. It is, indeed, quite obvious that a State possesses a legal interest in the protection of its territory from any form of external harmful action, as well as in the defence of the well-being of its population and in the protection of national integrity and independence. It would indeed be positively absurd to suggest otherwise. If a State did not possess a legal interest in such matters, how could Portugal have brought the Nauilaa case against Germany . . . ; how could Albania have brought against the United Kingdom in the Corfu Channel case . . . the claim arising out of the sweeping of mines in Albanian territorial waters? The point does not require elaboration.72

At least from the Australian perspective, even unintentional effects manifesting on its territory sufficed to breach territorial inviolability.

The International Court of Justice again faced the issue of territorial sovereignty in its 1986 Nicaragua judgment.73 The case involved Nicaragua’s assertion that the United States had breached its obligations under “general and customary international law” and “violated and is violating the sovereignty of Nicaragua” by “armed attacks against Nicaragua by air, land and sea”; “incursions into Nicaraguan territorial waters”; “aerial trespass into Nicaraguan airspace”; and “efforts by direct and indirect means to coerce and intimidate the Government of Nicaragua.”74

When considering these claims, the court acknowledged linkage between State sovereignty and the prohibitions of the use of force and coercive intervention, but unambiguously differentiated between them, noting that a single act may violate more than one of the prescriptive norms:

The effects of the principle of respect for territorial sovereignty inevitably overlap with those of the principles of the prohibition of the use of force and of nonintervention. Thus the assistance to the contras, as well as the direct attacks on Nicaraguan ports, oil installations, etc., . . . not only amount to an unlawful use of force, but also constitute infringements of the territorial sovereignty of Nicaragua, and incursions into its territorial and internal waters. Similarly, the mining operations in the Nicaraguan ports not only constitute breaches of the principle of the nonuse of force, but also affect Nicaragua’s sovereignty over certain maritime expanses. The Court has in fact found that these operations were carried on in Nicaragua’s territorial or internal waters or both . . . and accordingly they constitute a violation of Nicaragua’s sovereignty. The principle of respect for territorial sovereignty is also directly infringed by the

72. Id. ¶ 456.
74. Id. ¶ 250.
unauthorized overflight of a State’s territory by aircraft belonging to or under the control of the government of another State.\textsuperscript{75}

In the opinion of the court, then, territorial sovereignty enjoys independent valence. Indeed, it felt obligated to apprise the facts based on the “duty of every State to respect the territorial sovereignty of others.”\textsuperscript{76} Ultimately, the court found that the United States, through various actions, breached obligations under customary law with respect to intervention, use of force, and violation of territorial sovereignty.\textsuperscript{77} In doing so, it treated violation of territorial sovereignty as a self-standing primary norm with no less normative force than the other two.

In 2015, the International Court of Justice examined Costa Rica’s allegations that Nicaragua had sent armed forces into Costa Rican territory and dug a channel thereon, and Nicaragua’s contentions that Costa Rica had built a road in the contested area and caused transboundary environmental damage to Nicaragua.\textsuperscript{78} Both sides claimed that these actions violated their respective sovereignties. They disputed their opponent’s claims on the basis that no violation had occurred because the other side did not enjoy sovereignty over the areas in question. Extracts from the judgment exemplify the legal argumentation of the two States:

Costa Rica alleges that Nicaragua violated its territorial sovereignty in the area of Isla Portillos in particular by excavating in 2010 a \textit{caño} with the aim of connecting the San Juan River with the Harbor Head Lagoon and laying claim to Costa Rican territory. According to Costa Rica, this violation of sovereignty was exacerbated by Nicaragua’s establishment of a military presence in the area and by its excavation in 2013 of two other \textit{caños} located near the northern tip of Isla Portillos.\textsuperscript{79}

Nicaragua does not contest that it dredged the three \textit{caños}, but maintains that “Nicaragua enjoys full sovereignty over the \textit{caño} joining Harbor Head Lagoon with the San Juan River proper, the right bank of which constitutes the land boundary as established by the 1858 Treaty . . . .” Nicaragua further submits that “Costa Rica is under an obligation to respect the sovereignty and territorial integrity of

\textsuperscript{75} Id. ¶ 251.
\textsuperscript{76} Id. ¶ 213.
\textsuperscript{77} Id. ¶ 292.
\textsuperscript{79} Id. ¶ 66.
Nicaragua, within the boundaries delimited by the 1858 Treaty of Limits.\textsuperscript{80}

For its part, the court adopted the same territorial sovereignty-based line of analysis. As an example, it observed, “[s]ince it is uncontested that Nicaragua conducted certain activities in the disputed territory, it is necessary, in order to establish whether there was a breach of Costa Rica’s territorial sovereignty, to determine which State has sovereignty over that territory.”\textsuperscript{81} After answering that question, it unanimously found that “by excavating three caños and establishing a military presence on Costa Rican territory, Nicaragua has violated the territorial sovereignty of Costa Rica.”\textsuperscript{82}

In fact, the court left no room for debate regarding whether sovereignty can be violated as a matter of international law; it employed classic terms and concepts from the law of State responsibility, including “breach,” “responsible for breach,” and “obligation to make reparation,” thereby affirming that the obligation to respect territorial sovereignty is legally binding.\textsuperscript{83} Moreover, because the court found that Nicaragua had violated Costa Rica’s sovereignty, it held that it did not have to determine whether Nicaragua’s conduct amounted to a breach of the prohibition on the threat or use of force under the U.N. Charter or the Charter of the Organization of American States.\textsuperscript{84} Finally, the court also noted that its determination that Nicaragua had breached the territorial sovereignty of Costa Rica “provides adequate satisfaction for the nonmaterial injury suffered on this account.”\textsuperscript{85}

At no time in the case did either party assert the absence of a primary rule prohibiting violations of sovereignty. On the contrary, that rule lay at the heart of both sides’ claims. Nor did the court consider that option. All involved took the rule’s existence as a normative given, and the court rendered its judgment on that basis.

B. State Practice and Opinio Juris

Unlike judicial decisions, State practice and expressions of opinio juris are obligatory elements of any claim that an obligation to respect sovereignty is legally binding in customary international law. In this regard, it must be noted that States sometimes act in ways that affect, but do not violate, the exercise of sovereign rights of other States, such as imposing sanctions that

\textsuperscript{80} Id. ¶ 68.
\textsuperscript{81} Id. ¶ 69.
\textsuperscript{82} Id. ¶ 229.
\textsuperscript{83} “These activities were in breach of Costa Rica’s territorial sovereignty. Nicaragua is responsible for these breaches and consequently incurs the obligation to make reparation for the damage caused by its unlawful activities . . . .” Id. ¶ 93.
\textsuperscript{84} Id. ¶¶ 96–99.
\textsuperscript{85} Id. ¶ 139.
impact another State’s domestic economic activities. Additionally, the term “sovereignty” frequently appears in political statements without necessarily carrying legal weight. Thus, it is essential to be sensitive to customary law’s formal components of State practice and *opinio juris* when examining what States do, how they react to actions by other States, and what their officials say publicly. The examples that follow have been carefully selected as illustrations of the way in which States treat the issue of sovereignty in international law, rather than as an international relations concept.

States have characterized a plethora of incidents as violations of their territorial sovereignty. It must be cautioned that some involved the armed forces and therefore may also have implicated the prohibitions of the use of force or coercive intervention. The fact that States at times chose to discuss an incident as a breach of their territorial inviolability when the actions might also have crossed the use-of-force or coercive-intervention thresholds demonstrates that States consider the former to be a primary rule distinct from other primary rules that are based in the principle of sovereignty.

Unconsented-to aerial intrusions have long been considered a violation of the subjacent State’s territorial sovereignty. Noteworthy in this regard is the incident involving the downing of an unarmed American U-2 reconnaissance aircraft by the Soviet Union and the capture of its pilot in 1960. The United States did not protest the shoot-down. This reaction contrasts sharply with U.S. condemnation of the downing of an RB-47 reconnaissance aircraft by Soviet fighters and the imprisonment of its crew the same year.

The difference can only be explained by virtue of the locations of the aircraft at the time of the shoot-downs, since both incidents involved military aircraft performing similar missions in the same year. In the case of the U-2, the aircraft was in Soviet national airspace, which both sides appeared to acknowledge was subject to Soviet sovereignty. By contrast, the RB-47 was flying in what the United States characterized as international airspace above the high seas. Accordingly, while the former involved a violation of

---

86. See André Beirlaen, *Economic Coercion and Justifying Circumstances*, 18 REVUE BELGE DE DROIT INT’L 57, 67–69 (1984) (discussing the line that demarcates economic sanctions that are acceptable under international law from those that are not).

87. On the salience of examining incidents in the identification of international law norms, see W. Michael Reisman, *International Incidents: Introduction to a New Genre in the Study of International Law*, 10 YALE J. INT’L L. 1, 3 (1984) (“The normative expectations that political analysts infer from events are the substance of much of contemporary international law.”).


89. See id. at 136 (describing the incident).


national airspace, and thereby the Soviet Union’s territorial sovereignty, the latter, at least in the U.S. view, did not.

Four decades later, in 2001, U.S. military personnel aboard an unarmed EP-3 reconnaissance aircraft were detained after making a forced landing on a Chinese island following a mid-air collision with a Chinese fighter. China protested the nonconsensual landing, claiming, in part, that the American aircraft had “entered China’s airspace without permission, [thereby] seriously violating China’s territorial sovereignty.” The United States responded that the aircraft had been outside Chinese national airspace at the time of the collision and only entered it once in distress. It argued that while “military aircraft normally require permission to enter the territorial airspace of another nation,” the wrongfulness of penetrating foreign airspace while in distress is precluded. The dispute in the case was not over the existence of a rule prohibiting unconsented-to entry into another State’s sovereign airspace, but rather the application of a circumstance precluding wrongfulness. Indeed, by relying on the notion of distress, it can only be concluded that the United States accepted that the action would, absent such a circumstance, have amounted to an internationally wrongful act.

The debate over counterterrorist drone strikes similarly have focused attention on respect for sovereignty and territorial integrity. Although drone operations implicate the prohibition on the use of force, States regularly characterize them as sovereignty violations. For instance, Pakistan has repeatedly taken the position that “drone strikes on its territory are counterproductive, contrary to international law, a violation of Pakistani sovereignty and territorial integrity, and should cease immediately.” Russia Foreign Minister Sergey Lavrov has echoed this position, stating, “It is not right to violate the sovereignty and integrity of any State. We fully support Pakistan’s stance.” As explained below, the U.S. justification for the strikes likewise is framed in the narrative of sovereignty.

Analogous incidents have taken place at sea. In March 2007, fifteen British military personnel from the HMS Cornwall were searching a merchant dhow in the Persian Gulf when they were captured and subsequently detained for nearly two weeks by Iranian Islamic Revolutionary
Guard forces. Each side claimed the other had acted unlawfully based on the location of the incident; the United Kingdom stated that its forces were in Iraqi territorial waters, whereas Iran asserted that they were operating in Iranian waters. An Iranian Foreign Ministry spokesman, for example, argued that the British forces were “violating the sovereign boundaries” of Iran at the time of their seizure. An investigation by the British Ministry of Defence concluded that a factor contributing to the incident was “[t]he absence of an internationally agreed delineation of Territorial Waters (TTW) and [Northern Arabian Gulf] water-space coordination measures between Iraq, Iran and Coalition Authorities.” The dispute was conducted in the vernacular of the violation of territorial sovereignty.

Nine years after the British–Iranian incident, the Iranian Islamic Revolutionary Guard Corps captured two U.S. Navy riverine craft with military personnel aboard after they mistakenly penetrated Iranian territorial waters. The Revolutionary Guards labeled the incident an “illegal entry into the Islamic Republic of Iran’s waters.” Following negotiations, the ten individuals were released and the boats returned. Far from criticizing Iran for its actions in seizing the crew, Secretary of State John Kerry thanked them for their cooperation.

Standing maritime territorial disputes regularly generate breach of territorial sovereignty claims. Most well known are those over South China Sea maritime boundaries, which are disputed by multiple countries in the region. The U.S. Navy conducts “Freedom of Navigation” (FON) operations in areas where it believes China has made excessive maritime claims, and China typically shadows the warships and warns them out of its purported territory. Such disputes even arise among close allies. For instance, in a


well-known 1985 incident, a Coast Guard icebreaker navigated through the Northwest Passage, which the United States claims is an international strait, without seeking Canadian permission. In response, Canada “granted permission” (despite the lack of a request to that effect) for the voyage and, although the two countries agreed to the presence of Canadian observers onboard, the United States still disputed the Canadian claim of sovereignty over the waters.

On land, the abduction of Adolph Eichmann is a classic case regarding territorial sovereignty. Eichmann had headed the Gestapo’s Section for Jewish Affairs and was responsible for implementation of the Final Solution. Following the war, he fled to Argentina. In May 1960, the Israeli Mossad abducted Eichmann from Argentina and brought him to Israel for trial in the District Court of Jerusalem.

Following the incident, but before trial, Argentina elevated the issue to the U.N. Security Council. In a letter to the Security Council, it submitted that “[t]he illicit and clandestine transfer of Eichmann from Argentine territory constitutes a flagrant violation of the Argentine State’s right of sovereignty.” After considering the matter, the Council adopted Resolution 138, in which it observed that the “violation of the sovereignty of a Member State is incompatible with the Charter of the United Nations,” and requested that the Israeli government make appropriate reparation for its actions. Israel and Argentina subsequently issued a joint communiqué stating that they viewed as settled “the incident which was caused through the action of citizens of Israel that has violated the basic rights of the State of Argentina.”

In dealing with the question of whether a covert abduction operation in another country without that country’s consent negated its jurisdiction, the

104. Id.
106. Id.
107. Id.
109. S.C. Res. 138, pmbl., ¶ 2 (June 23, 1960). The explicit reference to a “violation of the sovereignty of a Member State” appears in the resolution’s preamble, whereas the operative part cites “acts . . . which affect the sovereignty of a Member State.” Id. pmbl., ¶ 1. This should not be interpreted as if the Security Council may not necessarily have regarded Israel’s action as unlawful. On the contrary, because the Security Council directed Israel to provide reparation “in accordance with . . . the rules of international law,” it must have concluded that a violation of international law had occurred; otherwise, no obligation to provide reparation would have materialized. Id. ¶ 2.
110. CrimC (Jer) 40/61 Att’y-Gen. of the Gov’t of Isr. v. Eichmann, PM 5722 ¶ 50 (1962) (Isr.) (quoting the Joint Communiqué of Israel and Argentina, reprinted in Att’y-Gen. of the Gov’t of Isr. v. Eichmann, 36 I.L.R. 59 (Isr., Dist. Ct. of Jerusalem 1961)).
District Court did not question the position that disrespect for territorial sovereignty can constitute a violation of international law. Clearly operating from the premise that such activities can do so, it concluded:

[N]ow that the Governments of Argentina and Israel have issued their joint communique . . . to the effect that both governments have decided to view as liquidated the “incident” whereby the sovereignty of Argentina was violated, the Accused in this case can certainly retain no right to base himself on the “violated sovereignty” of the State of Argentina. The indictment in this case was presented after Argentina had forgiven Israel for that violation of her sovereignty, so that there no longer subsisted any violation of international law. In these circumstances, the Accused cannot presume to be speaking on behalf of Argentina and cannot claim rights which that sovereign state has waived.\textsuperscript{111}

That an extraterritorial exercise of enforcement jurisdiction amounts to a violation of sovereignty of the State in which it occurs is now well settled in international law.\textsuperscript{112}

An interesting incident concerning territorial sovereignty over both national airspace and land occurred in 1978, when a Soviet spacecraft with a nuclear reactor onboard, Cosmos 954, reentered the earth’s atmosphere into Canadian airspace.\textsuperscript{113} During reentry, the spacecraft disintegrated and debris was scattered across a wide swath of Canada. Canada claimed for compensation, both on the basis of the Convention on International Liability

\begin{flushleft}
\textsuperscript{111} Id. ¶ 44.
\textsuperscript{112} “A state’s law enforcement officers may exercise their functions in the territory of another state only with the consent of the other state, given by duly authorized officials of that state.” \textsc{Restatement (Third) of Foreign Relations} § 432(2) (\textsc{Am. Law Inst.} 1986). Professor Louis Henkin suggested that “[w]hen done without consent of the foreign government, abducting a person from a foreign country is a gross violation of international law and gross disrespect for a norm high in the opinion of mankind. It is a blatant violation of the territorial integrity of another state . . . .” Louis Henkin, \textit{A Decent Respect to the Opinions of Mankind}, 25 \textsc{John Marshall L. Rev.} 215, 231 (1992). The fact that a State’s unauthorized exercise of extraterritorial-enforcement jurisdiction amounts to a violation of the other State’s sovereignty is also acknowledged in \textsc{Tallinn Manual 2.0}. \textsc{Tallinn Manual 2.0, supra} note 12, at 19, 67 (noting “[t]he Experts agreed that a violation of sovereignty occurs whenever one State physically crosses into the territory or national airspace of another State without either its consent or another justification in international law . . . .” and stating “the exercise of enforcement jurisdiction on another State’s territory constitutes a violation of that State’s sovereignty . . . except when international law provides a specific allocation of authority to exercise enforcement jurisdiction extraterritorially or when the State in which it is to be exercised consents”). Similarly, the U.N. High Commissioner for Human Rights has accepted that an extraterritorial exercise of jurisdiction may violate another State’s sovereignty. Rep. of the Office of U.N. High Comm’r for Hum. Rts. on the Right to Privacy in the Digital Age, ¶ 34, U.N. Doc. A/HRC/27/37 (June 30, 2014).
\end{flushleft}
for Damage Caused by Space Objects and “general principles of international law.” The dispute was settled in 1981 by means of a protocol between Canada and the Soviet Union. Of particular relevance to the issue of territorial sovereignty was the approach taken by Canada in its Statement of Claim:

The intrusion of the Cosmos 954 satellite into Canada’s air space and the deposit on Canadian territory of hazardous radioactive debris from the satellite constitutes a violation of Canada’s sovereignty. This violation is established by the mere fact of the trespass of the satellite, the harmful consequences of this intrusion, being the damage caused to Canada by the presence of hazardous radioactive debris and the interference with the sovereign right of Canada to determine the acts that will be performed on its territory. International precedents recognize that a violation of sovereignty gives rise to an obligation to pay compensation.

Regarding opinio juris, senior government officials in many nations have referred for decades to the violation of sovereignty in a fashion that qualifies as such. Soviet Prime Minister Khrushchev, for example, in pointing to the notion of coexistence, stated in 1959 that:

Apart from the commitment to nonaggression, [coexistence] also presupposes an obligation on the part of all states to desist from violating each other’s territorial integrity and sovereignty in any form and under any pretext whatsoever. The principle of peaceful coexistence signifies a renunciation of interference in the internal affairs of other countries with the object of altering their system of government or mode of life or for any other motives.

Note that Khrushchev not only confirmed Soviet acceptance of a rule prohibiting violation of territorial sovereignty, but also treated it separately from interference in internal affairs (coercive intervention).

Similarly, U.S. government representatives regularly offer expressions of opinio juris that operate from the premise of territorial sovereignty’s inviolability. To illustrate, numerous statements, including ones issued with other States, were made on this basis during, and in the aftermath of, the conflict between Georgia and Russia in 2009. Following the ceasefire, for example, the State Department’s spokesperson noted that Russia’s plans to build up its military presence in the Georgian regions of Abkhazia and South

Ossetia would not only breach the ceasefire agreement but also violate Georgia’s sovereignty and territorial integrity.  

More recently, Russian activities with respect to the Ukraine conflict, including Russia’s belligerent occupation of the Crimean peninsula since 2014, have consistently been portrayed as violations of sovereignty. President Obama characterized Russian actions as such when discussing the matter with President Putin in March 2014. The same month, the United States delivered a statement at the U.N. Human Rights Council on behalf of forty-two nations expressing concern over Russia’s “ongoing violation of Ukraine’s sovereignty and territorial integrity”, the G-7 did likewise. President Obama then stated that Russia “flagrantly violated the sovereignty and territory of an independent European nation, Ukraine” during his “Address to the People of Europe” in April, a claim he repeated at the NATO Warsaw Summit the same year.

Many relevant statements have been made with respect to counterterrorist operations. In a speech at National Defense University in 2013, President Obama noted that “our actions are bound by consultations with partners, and respect for state sovereignty.” Other members of his administration repeatedly made the same point. Attorney General Eric

Holder, speaking at Northwestern University School of Law, earlier had confirmed that “[i]nternational legal principles, including respect for another nation’s sovereignty, constrain our ability to act unilaterally.”127 His comments were especially salient, for the Justice Department renders the final decision on questions of law for the Executive Branch.128 The thread running through all of the statements has been recognition of an affirmative legal duty to respect the territorial sovereignty of other States in the conduct of U.S. counterterrorist operations; as a legal obligation, the duty represents a substantive rule, not simply the articulation of a broad normative principle or a restatement of the prohibitions of the use of force or coercive intervention.

Increasingly, senior U.S. government officials have acknowledged this duty with respect to activities in cyberspace. State Department Legal Adviser Harold Koh offered the first major statement on the matter in 2012 at an interagency legal conference convened at U.S. Cyber Command.129 In the speech, he addressed the issue of sovereignty head on.

States conducting activities in cyberspace must take into account the sovereignty of other States, including outside the context of armed conflict. The physical infrastructure that supports the internet and cyber activities is generally located in sovereign territory and subject to the jurisdiction of the territorial State. Because of the interconnected, interoperable nature of cyberspace, operations targeting networked information infrastructures in one country may create effects in another country. Whenever a State contemplates conducting activities in cyberspace, the sovereignty of other States needs to be considered.130

The position was clear. Remote cyber operations that cause effects in other States implicate, inter alia, the territorial sovereignty of those States. Koh spoke to the fact that it is incumbent on the State planning a remote cyber operation to consider whether the effects generated abroad breach the obligation to respect other States’ territorial sovereignty; while he did not

---


answer the question of when cyber operations violate sovereignty, he clearly accepted that in certain circumstances they do.

In a 2016 address at Berkeley Law School, Koh’s successor, Brian Egan, explicitly confirmed this point.

The very design of the Internet may lead to some encroachment on other sovereign jurisdictions. Precisely when a nonconsensual cyber operation violates the sovereignty of another State is a question lawyers within the U.S. government continue to study carefully, and it is one that ultimately will be resolved through the practice and opinio juris of States.\(^{131}\)

There was no suggestion that either of the former State Department Legal Advisers believed that sovereignty-related internationally wrongful acts in cyberspace were limited to uses of force or coercive intervention. On the contrary, both acknowledged that the principle of sovereignty applies in the cyber context and, by virtue of its legally binding nature, has operational significance.

This position tracks that contained in the 2014 U.S. submission to the U.N. Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (GGE). Stressing the application of sovereignty rules to the extraterritorial causation of effects, it noted,

Most cyber activities undertaken by States and other actors fall below the threshold of the use of force and outside of the context of armed conflict. Such activities, however, do not take place in a legal vacuum. Instead, they are governed by, inter alia, international legal principles that pertain to State sovereignty, human rights, and State responsibility.

State sovereignty, among other long-standing international legal principles, must be taken into account in the conduct of activities in cyberspace, including outside of the context of armed conflict. Because of the interconnected, interoperable nature of cyberspace, operations targeting networked information infrastructures in one country can have effects in many countries around the world. Whenever a State contemplates conducting activities in cyberspace, the sovereignty of other States needs to be considered.\(^{132}\)

---


132. Applicability of International Law to Conflicts in Cyberspace, 2014 DIGEST OF UNITED STATES PRACTICE IN INTERNATIONAL LAW, ch. 18, § A(3)(b), at 737. Interestingly, the Department of Defense’s own Law of War Manual emphasizes the obligation in an armed conflict to respect the sovereignty of other States during cyber operations because “cyber operations targeting networked information infrastructures in one State may create effects in another State that is not a party to the
Other States also apply a substantive, vice foundational, rule of territorial inviolability to cyber activities, distinguishing it from separate relevant primary rules of international law. For instance, indicative of the Netherlands government’s views were the opening comments of the Foreign Minister Bert Koenders at the 2017 European launch of *Tallinn Manual 2.0.* In his speech, he noted that “we mustn’t be naive. Cyber operations against institutions, political parties and individuals underline why we need the international legal principles of sovereignty and nonintervention in the affairs of other states.” In light of the hostile cyber operations he cited, the Minister can only have attributed operational consequence to the principle of sovereignty, which he distinguished from nonintervention. He went on to emphasize that “[t]he Tallinn Manual provides guidance on the application of long-established legal principles in the cyber domain: sovereignty, nonintervention, due diligence, and state responsibility.” That guidance, as explained, attributes primary-rule significance to sovereignty, a point that could not have been lost on the Netherlands Ministry of Foreign Affairs.

C. Sovereignty in International Fora

Both the U.N. Security Council and the General Assembly have treated the violation of sovereignty as a primary rule. For example, the Security Council resolution cited above in the Eichmann case specifically referred to “violation of the sovereignty of a Member State.” But among resolutions by U.N. organs, the General Assembly’s 1970 Declaration on Friendly armed conflict.” U.S. DEP’T OF DEF., OFFICE OF GEN. COUNSEL, LAW OF WAR MANUAL 1019 (2016). Although framed in the context of neutrality, such an operation in an international armed conflict could breach the State’s obligations with respect to both territorial sovereignty and neutrality. During a noninternational armed conflict, only the former would be breached, as the law of neutrality applies only to international armed conflicts. On neutrality, see Convention Concerning the Rights and Duties of Neutral Powers and Persons in War on Land arts. 5, 10, 17, Oct. 18, 1907, 36 Stat. 2310 (discussing the nature of international neutrality) and Convention Concerning the Rights and Duties of Neutral Powers in Naval War arts. 1–12, Oct. 18, 1907, 36 Stat. 2415 (establishing protocols for neutrality at sea). In the cyber context, see TALLINN MANUAL 2.0, supra note 12, at 553 (explaining the relationship between neutrality and cyber warfare).


134. Id. See also the report by noted international law experts that was commissioned by the government of the Netherlands which found that “[i]nternational law is based on a strict prohibition of the use of force and a duty to respect the sovereignty and territorial inviolability of other states. These rights and duties are a two-way street . . . .” ADVISORY COUNCIL ON INT’L AFFAIRS AND THE ADVISORY COMM. ON ISSUES OF PUB. INT’L LAW, No. 77 AIV/No. 22, CAVV, CYBER WARFARE 22 (2011).

135. Koenders, supra note 133.

Relations is perhaps the most significant general pronouncement of law bearing on the existence of such a rule. 137

The resolution’s text is especially noteworthy because it represents an unusual consensus during the divisive Cold War. In the Declaration, the General Assembly reaffirms “the basic importance of sovereign equality and [stresses] that the purposes of the United Nations can be implemented only if States enjoy sovereign equality and comply fully with the requirements of this principle in their international relations.” 138 This carefully negotiated verbiage implies that there are certain State actions that are not in compliance with—that is, violate—the principle of sovereign equality. This can only be so if sovereignty is more than an underlying principle; it must have operative effect.

Sovereign equality is one of seven principles highlighted by the Declaration. As to the principle, the resolution observes: “In particular, sovereign equality includes the following elements: . . . (d) [t]he territorial integrity and political independence of the State are inviolable . . . ” 139 It is telling that the reference to territorial inviolability appears with regard to a principle, sovereign equality, that is set out separately from the principle requiring States to refrain from the threat or use of force against the territorial integrity or political independence of other States. 140 This being so, it can only be understood as applicable in its own right.

Treaty law sheds further light on the existence of a rule prohibiting violations of territorial sovereignty. The U.N. Convention Against Transnational Organized Crime, for example, provides,

States Parties shall carry out their obligations under this Convention in a manner consistent with the principles of sovereign equality and territorial integrity of States and that of nonintervention in the domestic affairs of other States.

. . . Nothing in this Convention entitles a State Party to undertake in the territory of another State the exercise of jurisdiction and performance of functions that are reserved exclusively for the authorities of that other State by its domestic law. 141

Note how the first subparagraph distinguishes an act implicating sovereign equality and territorial integrity from one involving prohibited intervention, while the second deals with functions that are reserved to

---

138. Id. pmbl.
139. Id. (under preamble section titled “The principle of sovereign equality of States”).
140. Id. (under preamble section titled “The principle of equal rights and self-determination of peoples”).
another State, which, as explained above, is an additional basis for finding that an act violates sovereignty.

Other treaties likewise acknowledge the inviolability of territory. For instance, the Rio Treaty refers to “the inviolability or the integrity of the territory or the sovereignty or political independence of any American State.”\(^\text{142}\) The Charter of the Organization of American States provides that “[t]he territory of a State is inviolable.”\(^\text{143}\) It also sets forth a collective security scheme that applies “[i]f the inviolability or the integrity of the territory or the sovereignty or political independence of any American State should be affected by . . . any . . . fact or situation that might endanger the peace of America.”\(^\text{144}\)

Statements on sovereignty in the context of cyber operations have also begun to appear in international fora. In its 2013 report, the U.N. GGE, composed of representatives from fifteen States, stated that “State sovereignty and international norms and principles that flow from sovereignty apply to State conduct of ICT-related activities, and to their jurisdiction over ICT infrastructure within their territory.”\(^\text{145}\) Note how the GGE differentiates State sovereignty from the norms and principles that derive from sovereignty, thereby indicating a distinction between them. Also significant is the GGE’s treatment of the applicability of sovereignty to cyber conduct in a way that distinguishes it from the mere exercise of jurisdiction over cyber activities.

The GGE’s 2015 report expanded on this distinction:

In their use of ICTs, States must observe, among other principles of international law, State sovereignty, sovereign equality, the settlement of disputes by peaceful means and nonintervention in the internal affairs of other States. Existing obligations under international law are applicable to State use of ICTs.\(^\text{146}\)

In other words, the GGE singled out the principle of State sovereignty, differentiating it from that of nonintervention. Moreover, the GGE did so in a paragraph that discusses the law that regulates State cyber operations, thereby accepting that the principle of sovereignty limits the “use” of cyber technologies vis-à-vis other States as a matter of international law.

Finally, in 2016, the heads of State of the Shanghai Cooperation Organization issued a joint declaration in which they “call[ed] on the international community to develop a peaceful, secure, fair and open


\(^\text{144}\) Id. art. 25.


\(^\text{146}\) U.N. GGE 2015 Report, supra note 3, ¶ 28(b).
information space based on the principles of cooperation and respect for national sovereignty and noninterference in the internal affairs of other countries.” 147 It is clear that the obligation to respect the sovereignty of other States enjoys wide recognition globally, including in the cyber context.

IV. Concluding Thoughts

Corn and Taylor worry that a rule requiring respect for territorial sovereignty would impede important operations necessary to national security. They warn,

If the view were adopted that sovereignty is a rule violated by any action illegal under the domestic law of a state, states seeking to disrupt distributed terrorist cyber infrastructure would be under an obligation to either seek Security Council authorization or the consent of the state in whose territory the infrastructure resides. The nature of cyber operations and capabilities often require high degrees of operational security and the flexibility to act with speed and agility. Operating through a consent model could in important cases surrender operational initiative to the terrorist adversary or render response options unworkable. 148

Their concern is misplaced, for they treat the rule prohibiting violation of territorial sovereignty as absolute. This badly misstates the view of those supporting its validity. The rule’s proponents are clear that it does not apply to every remotely conducted cyber operation into another State’s territory. Indeed, they are divided over those operations that do breach inviolability. 149 The assertion that the rule on sovereignty somehow would leave a State defenseless in the face of serious threats to national security is also counter-normative. International law provides a robust toolbox for a State wishing to respond to hostile cyber operations that includes retorsion, 150 countermeasures, 151 actions based on the plea of necessity, 152 and self-defense. 153

Moreover, the consequences of the absence of such a rule for States that are the target of hostile cyber operations would be unacceptable. Although

148. Corn & Taylor, supra note 16.
149. See TALLINN MANUAL 2.0, supra note 12, at 19–21, 23.
150. Acts that, albeit unfriendly, are lawful, such as economic sanctions. Id. at 112.
151. Articles on State Responsibility, supra note 11, art. 22; TALLINN MANUAL 2.0, supra note 12, at 111 (Rule 20).
152. Articles on State Responsibility, supra note 11, art. 25; TALLINN MANUAL 2.0, supra note 12, at 135 (Rule 26).
153. U.N. Charter art. 51; TALLINN MANUAL 2.0, supra note 12, at 339 (Rule 71).
the precise threshold at which a cyber operation constitutes a use of force is unsettled in international law, it is undisputed that an offending operation must reach a high degree of severity. By the Corn and Taylor approach, operations falling below that threshold would be governed solely by the prohibition on coercive intervention. Yet a cyber operation that either does not affect a State’s domaine réservé or that is not coercive would not be encompassed in the prohibition. As an example, consider a State’s disruptive cyber operations directed against commercial cyber infrastructure in another State intended to give the former’s own companies a competitive advantage. The operations would lie beyond the prohibition because such activities are generally not considered to fall within the domaine réservé. Also problematic is the fact that cyber operations that are merely malicious or vindictive lack the requisite element of coercion.

In law as in life, what one sees depends on where one stands. Corn and Taylor take the perspective of those charged with conducting cyber operations into other States to defend the United States or otherwise advance its national interests. Thus, it is unsurprising that, given the ease by which cyber operations cross borders and their increasing frequency and severity, they do not want the hands of the Department of Defense tied. But one must wonder whether government departments charged with the conduct of diplomacy or fashioning policy responses to hostile cyber operations will be amenable to forgoing the option of labeling other States’ hostile cyber operations as unlawful unless they cross the coercive-intervention or use-of-force thresholds, especially in light of the fact that the vast majority of the operations do not. Additionally, bringing down the normative firewall in the manner they propose would bar the taking of countermeasures in response to many hostile cyber operations because the operations would not qualify as internationally wrongful acts.154

States facing cyber threats, but lacking the cyber wherewithal of the United States, are likewise unlikely to countenance a legal regime that opens the gates wide to hostile cyber operations. It would leave them legally defenseless in the face of most such operations, and factually dependent on the United States or other cyber powers for assistance in responding to them. It is worth recalling that States enjoy sovereign equality; they all get a vote in the development and subsequent authoritative interpretation of international law. That the international community will accept the

154. Of course, by the Corn and Taylor approach, qualifying a cyber response as a countermeasure may not be necessary because under the scheme many of the responses themselves would not breach an obligation owed to the other State. Yet, because responses need not be in-kind to qualify as lawful countermeasures, their approach would also remove the option of engaging in noncyber countermeasures. TALLINN MANUAL 2.0, supra note 12, at 128. The nonavailability of countermeasures might be especially problematic from a policy perspective as the United States continues to search for effective means by which to deter other States’ hostile cyber operations directed against it.
possibility of a cyber “wild west” below the intervention threshold is highly unlikely.

As has been demonstrated, Corn’s and Taylor’s arguments fly in the face of long-standing State practice, *opinio juris*, and judicial decisions as to the application of the primary rule of sovereignty that safeguards territorial integrity and inviolability. Indeed, they have cleverly attempted to shift the burden of persuasion in this regard. However, the evidence of the rule is so dense that those asserting its nonapplicability to cyber operations manifesting on the territory of another State must, as a matter of law, bear the burden of establishing why it does not apply to cyber operations. This they have failed to do. Instead, policy arguments and analysis are offered in the attempt to rebut a well-established legal notion.

Ultimately, Corn and Taylor conclude,

[W]hether and precisely when non-consensual cyber operations below the threshold of a prohibited intervention violate international law is a question that must be resolved through the practice and *opinio juris* of states, developed over time and in response to the need of states effectively to defend themselves and provide security for their citizens.\(^\text{155}\)

They are correct, but off course. Practice and *opinio juris*—and perhaps treaty law—will not determine whether territorial sovereignty is inviolable; it clearly is. Rather, practice and *opinio juris* will inform the contours of the rule as applied in the cyber context. Over time, it may even contribute to the emergence of *lex specialis* rules that provide for exceptions to the *lex generalis* rule protecting territorial integrity and inviolability. But for the present, such possibilities amount to nothing more than *lex ferenda*.

\(^{155}\) Corn & Taylor, *supra* note 16.
Spying and Fighting in Cyberspace: What is Which?

Gary Brown*

INTRODUCTION

Traditionally, espionage has inhabited a niche between order and chaos. States have recognized the existence of espionage and enacted domestic legislation to prohibit it, but international law is silent on the subject.¹ On the other hand, States accept espionage as part of the business of international relations and are generally tolerant of it. That may be changing, however. Cyberspace, especially the Internet, has become an integral part of everyday life. The use of cyberspace for espionage has generated difficult discussions about the nature of cyberspace, the extent of national sovereignty, and the importance of individual privacy, among other issues, all of which are relevant in a conversation about espionage. This article focuses on another issue, which is the overlap of espionage and aggressive cyber operations. Confusion about the intent behind an intrusion could lead to a misreading of aggressive intent, unnecessary escalation of tensions, or a false sense of security in the opening act of significant cyber aggression. This article also discusses the United States’ stance on dividing espionage into categories depending on the purpose.

Rapid improvements in computer technology and techniques, as well as the exponential rise in the amount of data stored online, have driven a closer look at the subject of cyber

---

* Gary Brown is a professor of Cyber Security at Marine Corps University.

¹ It could be cited as an exception that the International Court of Justice directed Australia to refrain from interfering with communications between Timor-Leste and legal advisers regarding current and future legal actions. See Press Release, Int’l Court of Justice, Questions relating to the Seizure and Detention of Certain Documents and Data (Timor-Leste v. Australia), Int’l Court of Justice (Mar. 3, 2014), http://www.icj-cij.org/docket/files/156/18076.pdf. That case stands as a solitary assertion, however, and applies to the special relationship between counsel and client, making its value as precedent questionable.
Espionage, in particular how it differs from traditional methods of spying. The speed of access and exfiltration in cyber espionage operations can rapidly result in libraries of information, dwarfing the information that can be obtained through more traditional methods of espionage. Although some of the issues discussed here are also relevant in traditional espionage operations, they have seemed less so in the past. They may have come to the forefront now because of the effectiveness and pervasiveness of cyber espionage. This article will focus only on cyber methods of espionage.

The United States defines espionage as “[t]he act of obtaining, delivering, transmitting, communicating, or receiving information about the national defense with an intent, or reason to believe, that the information may be used to the injury of the United States or to the advantage of any foreign nation.”

The distinction between cyber espionage and more aggressive cyber operations is critical under international law. Espionage has been considered unregulated under the international legal system – meaning cyber activities that constitute espionage are neither lawful nor unlawful under international law. As a result, States freely engage in espionage and generally accept it from other States, with results limited to punishing spies under domestic law and the expulsion of spies.

---


4 Whether or not espionage is prohibited by international law does not affect whether it may be prohibited or otherwise regulated domestically.
I. NOT ALL ESPIONAGE IS EQUAL

Historically, the United States appears to have agreed that international law should not apply to traditional espionage and that instead the punishment of spies should be left to domestic law. With the rise of cyber espionage, however, the United States has begun to change its position.6 “Traditional espionage encompasses a government’s efforts to acquire clandestinely classified or otherwise protected information from a foreign government,” explains cyber security expert, David P. Fidler. “Economic espionage involves a State’s attempts to acquire covertly trade secrets held by foreign private enterprises.”7 The United States manifested this distinction in the unprecedented indictment of five Chinese military officers for engaging in cyber espionage from China, in Administration statements critical of economic espionage, and in the U.S.-China agreement prohibiting cyber economic espionage for commercial gain, but is silent on other categories of espionage.8

---

In February 2013, the cyber security company Mandiant published a compelling portfolio of evidence tying the Chinese military to cyber economic espionage. That Mandiant chose commercial espionage for its deep-dive investigation appears to reflect the U.S. position that “economic espionage” should be treated differently than more traditional or “national security espionage.”

The United States treats as traditional espionage the theft of information more directly relevant to national security. U.S. concern over cyber espionage was reflected by then-National Security Agency Director, General Keith Alexander when he said “the loss of industrial information and intellectual property through cyber espionage constitutes the ‘greatest transfer of wealth in history.’” Although General Alexander’s statement has been criticized as exaggerated, there does appear to be a large, on-going transfer of possession of intellectual property through cyber-enabled espionage.

If espionage is to be split into two distinct categories, it may seem counterintuitive that economic espionage would be the more disfavored category. After all, economic espionage merely transfers net wealth and marginally decreases the incentive to innovate. It might make sense to treat economic espionage less seriously than traditional espionage, as the latter could have more significant national security implications.
directly and negatively affect national security. The United States has decided the opposite is true, perhaps because espionage directly benefiting national security is considered to have a longer, more established tradition. In addition, national security espionage may have come to be tolerated among States because it distributes knowledge that may increase the collective security of the community of nations by reducing surprise, increasing knowledge of intentions, etc.

In any event, there has been no clear international consensus that singles out economic espionage for denunciation. Currently, State responses to economic espionage include official condemnation, responsive sanctions or the use of other international tools to dissuade economic espionage. None of these indicate that it is treated differently than national security espionage.

Even if there were a concerted international movement to recognize the distinction between “good” and “bad” espionage, the details, at least to some degree, would be challenging. National security is a broad concept. It includes not just military forces, but also political stability – and the strength of the economy. Rational arguments can be made for a vast array of technologies contributing to “national security.” For example, energy technologies can benefit the military, food technology can increase a State’s self-sufficiency, and entertainment technology can increase the effectiveness of propaganda. The Commentary to Additional Protocol I notes that all information has some relevance for national security, and this is especially relevant with regard to cyber espionage.

---

13 It is too early to tell whether the U.S.-China agreement signals a change in the general international approach to the issue.
15 See Claude Pillot et al., Int’l Comm. of the Red Cross, Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949 566 (Yves Sandoz et al. eds., 1987).
II. ARE WE UNDER ATTACK?

Although the United States is engaged on the issue of categories of espionage, it has said little about the challenge of distinguishing between identical cyber activities undertaken for fundamentally different purposes. For instance, will virtual presence on a cyber system, without more information, be treated as espionage, remaining essentially unregulated, or be treated as preparation for cyber warfare akin to penetrating sovereign airspace with armed fighters or massing armed forces on the border?

In the purely physical world it is usually simple to distinguish espionage from bellicose activity. The weapons used to fight a war are generally distinguishable from those used to spy, both in nature and in quantity. For example, if a spy is armed at all it is likely with a sidearm or other light weapon. Spies usually work alone or in small groups. Basically, traditional spies look like ordinary citizens, or at most like ordinary criminals. It is often the intent of spies to look like insiders, or people who have permission to be where they are. Troops planning to engage in combat, on the other hand, appear to be what they are – combatants. Combatants are required to wear uniforms and carry their arms openly. They are normally armed with heavier weapons and present in larger numbers. These facts, together with the location of the individuals involved, generally make a determination of whether a particular activity is espionage relatively straightforward in the physical world.

Some cyber attacks are easy to define. For example, gaining access to a computer network and using the access to physically destroy attached computers or equipment is a cyber

---

16 Camouflage is a kind of “deception” perhaps, but the deconstruction of “cyber camouflage” I’ll leave to someone else.
attack. In more subtle cases, however, it can be difficult for the party on the receiving end of a cyber operation to distinguish between espionage and military attack (including actions leading up to an attack). Most cyber operations of any type require gaining unauthorized or secret access to an information system.\textsuperscript{17} When victims discover their cyber systems have been penetrated, determining what happened and whether information has been stolen or modified may not be easy if the attacker is patient and careful. It is often not immediately apparent whether the unauthorized access is intended for spying, for disruptive and destructive activities, or both. The potential damage is not limited to a physical location, as in the case of a saboteur, which ups the ante for cyber operations. To complicate the situation even more, the initial access may be for reconnaissance in advance of attack, so that the compromise and theft of data are preludes to future offensive operations. Finally, even if the initial purpose were espionage, access itself may embolden the hacker to commit a future attack.

Both espionage and warfighting benefit from acquiring access to as many systems as possible, to maximize either information gathering or the effect of a future attack. Given the nature of cyberspace, that might mean thousands of systems for either type of operation. Accordingly, both quantitatively and qualitatively, espionage and warfighting in cyberspace can be indistinguishable until the denouement.

Although merely gaining access to a network or computer is not a wrongful use of force or an armed attack under international law, the \textit{method} used might be.\textsuperscript{18} Some cases are simple. Invading a military base located across a national border, causing hundreds of casualties, for the purpose of seizing a hard drive containing sensitive information is not espionage – even if that is

\textsuperscript{17} Herbert Lin discusses these different actions in \textit{Offensive Cyber Operations and the Use of Force}, 4 J. NAT’L SECURITY L. & POL’Y 63, 64 (2010).

\textsuperscript{18} See \textit{TALLINN MANUAL ON THE INTERNATIONAL LAW APPLICABLE TO CYBER WARFARE} 195 (Michael N. Schmitt ed., 2013). The concepts of use of force and armed attack are from the U.N. Charter art. 2 \textit{§}4, 51.
the sole purpose of the excursion. It is a military attack. More subtle examples can be difficult
to parse. To facilitate espionage, a State might covertly dispatch a small military unit to break
into a secure facility for the purpose of inserting a flash drive into a network to upload malware
that will enable the collection of information. The smaller the unit, and the less force used, the
greater the likelihood the action will be seen as espionage – but at some point, such endeavors
constitute a significant breach of sovereignty or a wrongful use of force in violation of
international law.

Similarly, cyber activities undertaken for the purpose of collecting intelligence might
look like cyber attacks. The U.S. National Research Council has observed that there may be
situations where “the distinction between a cyberattack and [cyber intelligence gathering] may be
very hard to draw from a technical standpoint, since both start with taking advantage of a
vulnerability.” Both offensive cyber activity and cyber espionage rely on acquiring
unauthorized access to a system, and that often involves damaging a system in some way. The
damage may be reducing the effectiveness of the target system’s anti-virus software, decreasing
the effectiveness of its encryption programs, installing a back door or altering its operating
system, for example. If damage is defined to include activities that decrease effectiveness or
cause a system to cease its intended function, then each of these is an illustration of damaging the
targeted system.

The overlap of espionage and offensive operations in cyberspace appears to have been
recognized and has been addressed through policy and doctrinal definitions in the United States.
Cyber espionage is referred to as “computer network exploitation,” which is defined as “enabling

19 TECHNOLOGY, POLICY, LAW, AND ETHICS REGARDING U.S. ACQUISITION AND USE OF CYBERATTACK
CAPABILITIES 261 (William A. Owens et al. eds., 2009).
20 This concept of damage is also discussed below. See discussion infra Section III.D.
operations and intelligence collection capabilities conducted through the use of computer networks to gather data from target or adversary automated information systems or networks.”21 The critical phrase is “enabling operations,” which includes cyber activity that would otherwise be considered a cyber attack as noted above. That is, an enabling operation could logically include physically damaging one system to facilitate the gathering of intelligence from another system.

“Enabling” is distinct from the collection of intelligence; it is rather those things that permit the collection. As discussed above, these could include anything from a physical presence in a foreign computer center to damaging systems to make them exploitable. Of course, it also includes collateral actions necessary to collect intelligence, such as forcing a computer reboot to install malware or sending a phishing email, which are not, standing alone, the collection of intelligence. Some of these collateral activities are cyber attacks, but they are defined as part of an intelligence operation. This is a definitional overlap between two fundamentally different categories of activity.

Occupying the space between cyber espionage and cyber aggression is Operational Preparation of the Environment (OPE). The Department of Defense defines OPE as “[t]he conduct of activities in likely or potential areas of operations to prepare and shape the operational environment.”22 OPE could include cyber operations to penetrate systems, introduce malware or undertake other actions in preparation for offensive action. These activities occur in the absence of armed conflict, although conflict may be anticipated.

Pre-positioning cyber capabilities on networks or computer systems, by itself, does not constitute cyber aggression, and is not quite espionage, because it is not collecting intelligence. This activity is rather some unique category falling between espionage and attack. Although capabilities are prepositioned in the kinetic world as well, the legal issues are easier to deal with in the physical world. For example, there is little doubt that concealing a weapons cache in another State’s territory is preparation for armed attack. On the other hand, obtaining access to a system often fails to signal what kind of follow-on action is anticipated. This ambiguity is one thing that makes cyber operations uniquely challenging.

Similarly, many pre-positioned capabilities provide the ability to engage in either espionage or aggressive activity, and so acting to emplace these capabilities may be mistaken for either of the other two. For example, malware that allows its controller to log on a system with administrator privileges would provide the opportunity to view or copy information on a network, as well as delete information and take other actions that could physically damage the system, i.e., constitute an attack. Obtaining and maintaining this kind of pre-positioned capability could be seen as the equivalent of planting explosives to be used at a future point.

This article will not address cyber OPE as a unique category. Although there are doctrinal and policy reasons for treating it as distinct, OPE can be included in this discussion by looking at it as an intelligence activity that has the potential to be mistaken for aggression.

III. A FRAMEWORK FOR ANALYSIS

There are more commonalities than distinctions between cyber espionage and cyber aggression. The framework below provides a broad overview of the steps involved in cyber
operations, followed by brief vignettes drawn from actual events that apply the framework. This analysis helps delineate the gray areas between cyber espionage and other cyber operations.

Put simply, any cyber operation requires identification, penetration, presence, exploitation and harm. I illustrate this using a pretend state-sponsored hacker named P0wn$z.

The first requirement for any operation is determining the target. The identification of a cyber system is the least elegant step. P0wn$z might do this by using a bot to conduct a massive survey of cyber systems, seeking out those with typical characteristics for the system he wants to target; for example, some SCADA systems have characteristics that make them easy to spot on the Internet. P0wn$z will be looking for the type of systems he wants that have vulnerabilities, such as unpatched software or unchanged default passwords. In this way, P0wn$z can build an extensive database of potential targets that he can sell to the highest bidder or use for his own purposes.

Once P0wn$z finds the system he wants to target, initial penetration of a system can be accomplished in a variety of ways. For Stuxnet, the cyber operation that destroyed nuclear centrifuges in Iran, it was through a worm. In the case of Operation Buckshot Yankee, it was most likely effected by the strategic placement of flash drives containing malware that were eventually used on official systems. Many system penetrations use the tried and true method of phishing emails, which are often cleverly crafted using information available from social media.

23 ICS-CERT noted the ease of identifying some of these systems in Dep’t of Homeland Sec., Incident Response Activity, ICS-CERT MONITOR, Jan.-Apr. 2014, at 1, 2.
26 See III.B. below.
Regardless of the method, the purpose is to gain and elevate access to the target system. That is, the goal is to get on the system and ideally to gain credentials as a system administrator.

After gaining access, the next thing P0wn$z wants to do is establish a persistent presence on the system. Operating systems and anti-virus software may be updated and passwords may change, for example. P0wn$z wants to access the system repeatedly. To exfiltrate large amounts of data, P0wn$z will spread the downloads over the course of several days or weeks to avoid being noticed by network monitoring tools. Besides, new information will be added to the system constantly, and a persistent access may yield results for many years. To establish persistent access, P0wn$z may install additional malware or create additional accounts on the system, for example, to provide a back door for future use.

The fourth step in the operation is exploitation of the access to gain information. As noted above, this may involve the exfiltration of information to a server located anywhere in the world, from where P0wn$z can move it later to where it will be analyzed. Exploitation might also involve real time monitoring of email content or system usage data to get inside the decision loop of the target organization. Another use of exploitation is to gather system information so that the system itself can be degraded or damaged.

Using the information to cause harm is the ultimate goal of a cyber operation, whether espionage or military. An espionage operation would seek to use the information gathered to do damage to the national security of the target State. In some cases, the target’s national security is weakened because a potential adversary has learned some strategic secret, such as where troops plan to strike, or a technical secret such as how to defeat a radar system. In some cases, the relative security of the victim State is reduced because a rival State has narrowed the victim’s lead in some strategic technology. In either case, the spying State benefits and the target State
suffers a detriment. It could be argued that no harm is intended or follows when “friends” spy on “friends,” as when the United States obtained access to the German Chancellor’s cellphone.\(^{27}\) The term “harm” as defined here includes changes in the relative advantage between States, because spying friends are potential future adversaries. As Henry Kissinger famously noted, “America has no permanent friends or enemies, only interests.”\(^{28}\)

As noted earlier, the United States sees a subset here. According to the United States’ view, using the pilfered information for commercial gain is fundamentally different from using it for the advancement of national security.\(^{29}\) China, however, has asserted that a State’s economy is an essential part of its national security, so damaging one State’s economy or benefiting the economy of another is the same as any other use of information obtained through espionage.\(^{30}\) Whether one position is preferable in law will not be discussed here. It can also be difficult to determine whether a particular operation is undertaken for the purpose of commercial gain or whether it incidentally results in commercial gain. This difficulty in distinguishing between the facts underlying the two positions is addressed in the scenarios below.

In more aggressive operations the harm intended might be actual damage to the host computer system, destruction of critical data, or damage to industrial systems connected to the network, for example. The important thing to note is that penetration, presence and exploitation


\(^{28}\) Kissinger was echoing a classic foreign policy position. This international reality is what made the 2010 revelation of the no spying agreement among the “Five Eyes” countries so surprising. Gordon Corera, *Spying Scandal: Will the ‘Five Eyes’ Club Open up?*, BBC (Oct. 29, 2013), http://www.bbc.com/news/world-europe-24715168.


\(^{30}\) In the end, there may be little difference between the United States and Chinese views on this matter, though the United States tends to phrase its position in terms of how the loss of information harms its national security rather than how obtaining it would improve its security. *See EXEC. OFFICE OF THE PRESIDENT, ADMIN. STRATEGY ON MITIGATING THE THEFT OF U.S. TRADE SECRETS* (2013).
may be precisely the same, whether the operation is intended for espionage or aggression. It is only with the harm that the two types of operation become distinguishable. This similarity throughout most of the operation creates challenges for legal and policy frameworks, as will be evident in the description of the operations below.

The examples below illustrate how penetration, presence, exploitation and harm apply in some publicly reported cyber operations. The crucial first step of identification is left for another paper, as it is focused on technology and intelligence collection rather than policy and law.

A. Undersea Cable Tapping

Cable tapping is discussed as a cyber operation because most Internet traffic passes through submarine cables. The United States has reportedly collected information from undersea communications cables for years. In the 1970’s the United States attached recording boxes to Soviet undersea cables. 31 Later, the United States (and others) may have tapped into submarine cables at repeater junctions under the sea.32 From published reports, this appears to be a blended cyber-kinetic method that introduces a new item of physical equipment to a system to collect cyber intelligence. An operation that collects such huge amounts of information is a gold mine of espionage. The penetration of the undersea cables that cumulatively carry 99% of the world’s Internet traffic was most likely accomplished through a variety of physical means.33 As espionage equipment was physically attached to the cables, it continued to maintain the presence

32 Id.
on the system. The exploitation was through a variety of means, as well, the most entertaining being the divers retrieving tapes from Soviet cables every few weeks.34

The complicating factor in this operation is the scale. If all the data moving through the cable is collected, it includes both national security and purely commercial data – and, of course, an enormous amount of personal information that raises constitutional issues beyond the scope of this article. The physical devices designed to be attached to undersea cables could include the capability to jam or otherwise interfere with electronic traffic passing through the cables. This would be an especially desirable way to deny communications during a conflict, because the system could be restored essentially cost-free after the conflict. Even in a case like this one that seems like simple espionage, the technology injects an element of doubt concerning the actor’s intentions. The mere presence on the system could be espionage or preparing for conflict.

B. Operation Buckshot Yankee (OBY)

In 2008, DoD’s classified military computer networks were compromised by malware. A flash drive pre-loaded with targeted malware was inserted into a military laptop at a base in the Middle East. The malicious code copied itself onto U.S. Central Command’s computer network, from where it spread across the military system, infecting both classified and unclassified computers. The purpose of the malware was to discover what information was available on the network, report back to its controller and then exfiltrate desired information. DoD concluded the malware was distributed by a foreign intelligence agency.35

34 Khazan, supra note 31.
Perhaps the most interesting feature of the malware used here was its ability to jump the air gap between the classified and unclassified computer systems, a capability critical to the success of the Stuxnet operation.\textsuperscript{36} When legitimate users used a flash drive to transfer information between systems, the malware was designed to ride the flash drive for the initial infection, and later to cause information to hitchhike on the drive from the classified to the unclassified system. From the unclassified system, sensitive information could be transferred over the Internet.\textsuperscript{37}

OBY was a straightforward cyber espionage operation. It appeared to target an official information system with the intent of gathering national security information to use for national security purposes. There were no reports that the malware used was capable of damaging the compromised system, so there was little chance of mistaking the intent of the spying State.

\section*{C. F-35 Plans}

Although few details have been released, in 2007 China hacked U.S. government contractor computer networks and obtained millions of pages of F-35 (also referred to as the Joint Strike Fighter or JSF) technical data.\textsuperscript{38} “According to a report from Independent Journalism Review, the U.S. Naval Institute speculates that the J-31 was ‘designed using

\begin{itemize}
\item \textsuperscript{36}“An air-gapped computer is one that is neither connected to the internet nor connected to other systems that are connected to the internet.” Kim Zetter, \textit{Hacker Lexicon: What Is an Air Gap?}, Wired (Dec. 8, 2014), http://www.wired.com/2014/12/hacker-lexicon-air-gap/.
\end{itemize}
This may at first appear to be another typical espionage case, and perhaps it is. It also helps illuminate the complexity of applying the U.S. position on good and bad espionage. U.S. officials noted that the theft of this data caused great damage to U.S. interests, giving away a substantial U.S. advantage in aviation, while reducing the lead time and costs to adversaries working to develop stealth technology themselves. The harm that resulted to the United States’ lead in stealth aircraft technology and the benefit to China’s program are typical of espionage operations. The pertinent distinction here is that the information was apparently given to a manufacturer, Shenyang Aircraft Corporation, which presumably profited from it, while improving China’s air force and national security.41 Where is the line between strategic technology and private sector technological advances? It may be difficult to draw. For example, solar power could make troop deployments more efficient by reducing fuel needs. Automobile technology may improve military vehicles. An advance in health sciences may improve battlefield medicine. Virtually any manufacturing technology can be related to national security.

D. Equation Group

This recently reported case is an example of supply chain exploitation. It simplifies the job of spying if the target’s hardware is manipulated in advance to permit unauthorized access. In this case, a State’s security service is reported to have installed capabilities on firmware

---

basically built-in software that controls the hardware) before it arrived at its destination. As reported, “[t]he malicious firmware created a secret storage vault that survived military-grade disk wiping and reformatting, making sensitive data stolen from victims available even after reformatting the drive and reinstalling the operating system.”

In this case, penetration and presence occur before the equipment becomes the target; exploitation is available as soon as it is worthwhile. Although this capability may not be able to damage the system directly, if you cannot use the targeted device as intended any more, but it still works, has there been an attack? If a system contains any sensitive information, once the penetration is discovered, the hardware is not usable. Functionally, it has been destroyed. Because of the time involved in an operation of this type, there is less risk of escalation, but there is still the question of characterization. Is it merely espionage when the process requires functionally destroying the target system? Once again, the scale of all things cyber may play a role. Destroying a few items in the name of espionage may mean little. What if a supply system penetration is discovered that affected hundreds of thousands of computer chips, routers or other components? At some point, it seems this could become something more than simply spying.

E. SCADA Systems

Utilities and modern manufacturing processes are often managed by computerized industrial control systems, most commonly referred to as Supervisory Control and Data Acquisition (SCADA) systems. SCADA systems are vital in the modern industrial world,

---

42 Dan Goodin, How “Omnipotent” Hackers Tied to NSA Hid for 14 Years – and Were Found At Last, ARS TECHNICA, (Feb. 16, 2015), http://arstechnica.com/security/2015/02/how-omnipotent-hackers-tied-to-the-nsa-hid-for-14-years-and-were-found-at-last/.

43 See id.

44 SCADA is the term most generally recognized in legal and policy discussions about cyber operations to describe computer systems that facilitate the remote control of industrial and utility systems, even when the systems might
controlling things as critical as drinking water plants, steel processing, auto manufacturing and electrical power grids. SCADA systems are designed for long lifespans and reliability, with security often considered a lower priority. They do not contain much information of interest, except to those who might be planning a cyber attack on the system. On the other hand, the lack of security on a networked SCADA system can make it an inviting target for hackers hoping to gain access to connected systems. For example, the massive breach of Target’s computer system appears to have been facilitated by computer credentials stolen from the company’s air conditioning service provider. That incident resulted in the exposure of 70 million Target customers’ personal data. Thieves and military planners may have good reasons for hacking into SCADA systems – but spies remain problematic.

Because States do not store secrets on utility systems, and the systems generally contain only information about the utilities themselves, any information that could be obtained from a SCADA system is probably only useful as reconnaissance for a future attack. Does it follow that merely establishing persistent presence on a SCADA system could be taken as aggressive? In most cases the intelligence value of any information is so low that analysts might assume the operation is not an exercise in simple espionage, but rather a prelude to aggression. U.S. SCADA systems are frequently targets of cyber operations. The potential harm is considerable. Espionage and operations with more aggressive intent seem particularly difficult

46 Id.
to distinguish in these cases. In 2014, a hacker caused “massive damage” to a steel plant in Germany. Just before the final step, it may have been impossible for an administrator of the steel plant’s systems, having discovered a hacker inside the system, to know whether the intruder was in the final stages of preparing for the destructive attack or merely spying, which creates a risk of miscalculation.

A final case that may help bring all the threads together is the 2014 Sony hack. In that incident, hackers gained access to Sony’s computer network. The hackers released a huge amount of business data, emails, personal data of employees, salary information, full copies of unreleased movies, and more. At some point the operation took a hostile turn and destroyed data on the servers.

The facts of the incident work well for this discussion if we speculate about a similar attack on FBI servers. In such a case, the FBI might detect the intruders at an early phase of the operation: while they are penetrating the federal computer system, establishing a persistent presence or exfiltrating sensitive anti-terrorism data, for example. At any of these times, it would appear to be nothing more than an espionage case. Then, perhaps without warning, the operation might turn aggressive. The same malware capabilities used to exfiltrate data might be used to delete (i.e., destroy) data and to render much more data inaccessible by corrupting the master boot records of hard drives. Would such a virtual destruction of a critical government information system rise to a level justifying a kinetic response? The United States acknowledged

51 Deleting the master boot record of a hard drive makes it practically impossible to access the data on the drive, even though it is still present.
the possibility that a cyber operation could justify actions in self-defense in its 2011
*International Strategy for Cyberspace*: “When warranted, the United States will respond to
hostile acts in cyberspace as we would to any other threat to our country.”52 If an apparent
espionage operation can so quickly turn destructive, at what point is a State justified in
aggressively acting in anticipation of a cyber attack?53

**CONCLUSION**

As discussed here, the tactics and techniques used in espionage and military operations in
cyberspace are often identical. Although when reviewing the results of cyber activity, it may be
easy to determine what the purpose of the action was, mid-operation – when responses are being
considered – there is great potential for international misunderstanding and miscalculation.
There is not an easy fix; it is simply a situation with which the international community must
contend. Espionage will continue to be required as part of a responsible strategy prior to military
action, and there is no indication the world’s “second oldest profession” will end even in the
absence of aggressive intent, because it supports economic and diplomatic strategies, as well.

The observation that for a significant duration of a continuous cyber event it is impossible
to distinguish between espionage, preparing the environment for a cyber attack, and the
beginning of a cyber attack is unlikely to change the behavior of States. Despite the potential
pitfalls set out here, States will continue to pursue courses of action – in this case the cyber

52 See Exec. Office of the President, Int’l Strategy for Cyberspace: Prosperity, Security, and Openness
in a Networked World 14 (2011),
53 Of course, the difficulty of attributing cyber actions to a particular State could mean that the target of aggressive
self-defense would be uncertain, but that just makes the situation more dangerous, as even uncertain national leaders
might feel compelled to “do something” to demonstrate to a restive population they are still in control.
options – they think best serve their own interests. Cyber espionage in particular is likely to continue to increase, as it results in the collection of huge amounts of strategic data for intelligence agencies. Rather than focusing on the unattainable, policy efforts would be better spent elsewhere. States should not attempt to create a different standard for cyberspace espionage, and for different types of espionage in cyberspace. Often, military operations in cyberspace and cyber espionage are distinguishable only by intent, which is difficult or impossible for the victim to ascertain. States should rather focus on the actual actions, as it is the behavior and the effects that determine international legality, not the intent of the actor. States might be reluctant to agree to stop engaging in strategically lucrative activity in return for increased international cooperation, but the expedient path of trying to divide cyber activities into categories of good and bad does not seem to have resulted in increased international understanding about state-sponsored cyber activities.

In a loosely governed environment like cyberspace, a shared understanding of the boundaries on acceptable behavior may be the best way to avoid unnecessary tension, or even escalation to hostilities. Discussions about what is okay and what is not would be easier if they focused purely on the activities themselves, rather than trying to pigeonhole cyber behaviors according to intent.
The Sony and DNC hacks exposed a significant gap in the international law of cyberspace: States are not being held accountable for the vast majority of their harmful cyberoperations, largely because classifications created in physical space do not map well onto the cyber domain. Most injurious and invasive cyberoperations are not mere “cybercrimes” and do not constitute “cyberwarfare,” and states appear unwilling to extend existing definitions of unlawful acts permitting countermeasures to such conduct (possibly to avoid creating precedent restricting their own cyberoperations). As a result, victim states have few effective and non-escalatory responsive options, and the harms associated with these incidents tend to lie where they fall.

This Article draws on tort law and international law principles to construct a comprehensive system of state accountability in cyberspace, where states are both liable for their harmful acts and responsible for their wrongful ones. Namely, it identifies international cybertorts—acts that employ, infect, or undermine the internet, a computer system, or a network and thereby cause significant transboundary harm—as distinct from cybercrime and cyberwarfare. Not only does this term distinguish a specific kind of harmful act, it highlights how the principle of state liability for transboundary harms (which holds states accountable for the harmful consequences of their lawful and unlawful activities) could usefully complement the existing law of state responsibility (which applies only to unlawful state acts) to create a comprehensive system of state accountability in cyberspace. Furthermore, imposing state liability for international cybertorts preserves a bounded grey zone for state experimentation in cyberspace, minimizes the likelihood that victim states will resort to escalatory responses, and increases the chance that those harmed by cyberoperations will be compensated.

* Executive Director, Information Society Project; Research Scholar and Lecturer in Law, Yale Law School. For productive conversations and useful insights, many thanks to Dapo Akande, BJ Ard, Jack Balkin, Jack Goldsmith, Claudia Haupt, Ido Kilovaty, Asaf Lubin, Torey McMurdo, Michael Schmitt, Beatrice Walton, Sean Watts, and Sheldon Welton. Earlier drafts were much improved from feedback associated with presentations at the ISP Resident Fellows Workshop, Andrew Chin’s Cyberspace Law course, and the Yale PhDs in Law tea. As always, many thanks to Douglas Bernstein for clarifying edits.
International Cybertorts (DRAFT)

Introduction .................................................................3
I. A Problem Without a Name .............................................8
   A. Modern International Law’s Limitations on Self-Help .............9
      1. Charter Restrictions on the Use of Force ......................10
      2. Customary Limits on the Use of Countermeasures ............11
   B. The Need for Effective, Non-Escalatory Deterrents .............13
      1. Practical Limits of Deterrence-by-Denial ....................13
      2. Practical and Legal Limits of Deterrence-by-Punishment ...15
      3. State Paralysis ..................................................19
II. State Liability for International Cybertorts ........................21
   A. A Distinct Kind of Harmful Cyberoperation .....................21
   B. International Cybertorts ..........................................24
      1. Relationship with Cybercrime and Cyberwarfare .............25
      2. Relationship with Data Destruction and Ransomware .........27
      3. Relationship with Cyber Exploitation and Cyberespionage ...29
   C. State Liability .....................................................31
      1. State Liability for Transboundary Harms .......................33
      2. Benefits of State Liability for International Cybertorts ......36
      3. State Liability in Cyberspace: Questions to Be Considered ...39
III. State Responsibility for Internationally Wrongful Acts .............47
   A. The Law of State Responsibility ..................................47
      1. Breach of an International Obligation ..........................47
      2. Attribution ......................................................48
      3. Reparations .....................................................50
   B. Cyber-Facilitated Interference ....................................51
      1. Unlawful Interference: Violations of State Sovereignty and
         Interventions ......................................................51
      2. An Elusive Line Between Lawful and Unlawful Interference
         ........................................................................53
      3. Increased Likelihood of Interference ............................56
   C. How State Liability Might Minimize Resort to Countermeasures .58
IV. A Comprehensive System of State Accountability in Cyberspace ......61
   A. State Interest in Developing the Law ..............................61
   B. Existing Implementation Mechanisms ...............................65
   C. A New Institution ....................................................66
   D. A Preferable Means of Law Creation ................................69
      1. The Unlikelihood of a Comprehensive Cybersecurity Treaty 69
      2. The Difficulty with Developing Customary International Cyber
         Law ........................................................................71
      3. The Benefits of Institutional Legal Development ...............72
Conclusion ........................................................................72
INTRODUCTION

In 2014, the North Korean “Guardians of Peace” hacker group raided Sony Entertainment Pictures servers and publicized extensive confidential data, including previously-unreleased films, executives’ embarrassing personal emails, actors’ passports and aliases, and Sony employees’ personal and medical information.\(^1\) In response, the United States took the then-unprecedented move of publicly attributing the Guardians’ cyberoperations directly to the state of North Korea and imposing new financial sanctions.\(^2\) Experts estimate that the costs of the Sony hack include $70 million in direct damages and more than $130 million in indirect damage (such as leaked trade secrets and lost revenue).\(^3\)

Eighteen months later, on the eve of the Democratic National Convention, WikiLeaks released approximately 18,000 emails written by top officials in the Democratic National Committee (DNC) that criticized and mocked then-presidential-hopeful Senator Bernie Sanders, sowing discord in an already-divided political party.\(^4\) A private cybersecurity firm determined that the emails had been obtained by hacker groups associated with the Russian government;\(^5\) months later, the Obama Administration formally attributed the DNC hack to Russia,\(^6\) sparking a debate about whether the hack and

---


5. Hamburger & Tumulty, supra note 4.

6. David E. Sanger & Charlie Savage, *U.S. Says Russia Directed Hacks to Influence Elections*, N.Y. TIMES, Oct. 7, 2016, at A1. While Russia has repeatedly denied having officially orchestrated the attacks, in May 2017 Russian President Vladimir Putin made the startling suggestion that “patriotically minded” private Russia hackers could have meddled
subsequent info dump altered the results of the 2016 presidential election. The economic expenses associated with the DNC hack were high; the political costs are impossible to calculate.

Predictably, the Sony hack and DNC hack were popularly termed “cyberwarfare”\(^7\)—and, equally predictably, these characterizations were followed by a spate of academics and specialists clarifying that they didn’t satisfy the legal requirements for that categorization.\(^8\) But if these weren’t cyberwarfare, what were they? They were cyberespionage and transnational cybercrime—but they were also something more. Unlike most cyberespionage, the stolen information was intentionally publicized with an apparent intent to cause harm. Unlike most transnational cybercrimes, the cyberoperations were state-sponsored—and while individuals in the Guardians of the Peace and Russian hacker groups can theoretically be held criminally liable, North Korea and Russia cannot.\(^9\)

\(^7\) With regard to the Sony hack, for example, former Speaker of the House Newt Gingrich tweeted, “No one should kid themselves. With the Sony collapse America has lost its first cyberwar. This is very very dangerous precedent.” @newtgingrich, Twitter (Dec. 17, 2014; 2:05 PM), https://twitter.com/newtgingrich/status/545339074975109122. Senator John McCain has described the DNC hack as an “act of war.” Theodore Schleifer & Deirdre Walsh, McCain: Russian Cyberintrusions An “Act of War”, CNN, Dec. 30, 2016, http://www.cnn.com/2016/12/30/politics/mccain-cyber-hearing/.


suggested that they might be internationally wrongful acts—violations of a state’s international obligations—but scholars are divided on that question.\(^{10}\) Importantly, the United States never claimed either hack was a violation of international law—instead, as evidenced by President Obama’s description of the Sony hack as “cybervandalism,”\(^{11}\) states are casting about for a way to characterize such acts negatively without explicitly labeling them as unlawful (and thereby setting a precedent that might limit their own cyberoperations). In short, despite being widely recognized as important and possibly even world-altering,\(^{12}\) there is no obviously accurate term for cyberoperations like the Sony and DNC hacks.\(^{13}\)

This Convention creates no international law crimes; rather, it creates international law obligations to enact domestic law and render mutual assistance. And, as with most criminal law regimes, this treaty is meant to govern the unlawful actions of individuals, not of states. All in all, it is wholly inadequate for addressing state-sponsored cyberoperations—notwithstanding the fact that, with a few notable (and contested) exceptions, the vast majority of significantly harmful and intrusive cyberoperations appear to have been sponsored by states. Furthermore, attempts to investigate and prosecute individuals for transnational cybercrimes under the treaty have not been markedly successful. Cf. Michael Glennon, State-level Cybersecurity, POL’Y REV. (Feb. 1, 2012), available at http://www.hoover.org/research/comparing-strategic-and-legal-features-cyberwar-drone-warfare-and-autonomous-weapon-systems.

\(^{10}\) The Sony hack might be considered a violation of U.S. sovereignty, insofar as there was manipulation of cyber infrastructure and the insertion of malware. Schmitt, supra note 8. It is less clear if the DNC hack could be similarly characterized, given that comprising computer systems and stealing data might be considered routine state practice in cyberspace. Sean Watts, International Law and Proposed U.S. Responses to the D.N.C. Hack, JUST SECURITY (Oct. 14, 2016; 8:48 AM), https://www.justsecurity.org/33558/international-law-proposed-u-s-responses-d-n-c-hack/. Neither cyberoperation would seem to be sufficiently coercive to meet the standard required for prohibited intervention.


\(^{13}\) The Sony hack has also been described as an act of cyberterrorism. See Ellen Nakashima, White House Says Sony Hack Is A Serious National Security Matter, WASH. POST, Dec. 18, 2014, https://www.washingtonpost.com/world/national-security/white-house-says-sony-hack-is-a-serious-national-security-matter/2014/12/18/01eb8324-86ea-11e4-b9b7-b8632ae73d25_story.html?utm_term=.a270fdd4a97e. However, as with its root
In addition to being difficult to classify, cyberoperations like the Sony and DNC hacks have distinct, often economic, harms. According to a 2014 Pricewaterhouse Coopers survey, the number of institutions reporting cyberoperations costing more than $20 million increased 92% from 2013 to 2014 (with an 86% increase in the number of institutions reporting attacks by nation-states). Experts estimate that malicious cyberoperations cost the U.S. economy between $120 and $167 billion in 2015 alone; others calculate that they will cost global businesses more than $6 trillion annually as of 2021.

The Sony hack, DNC hack, and other recent malicious cyberoperations have exposed a significant gap in the international law of cyberspace: states are not being held accountable for the vast majority of their harmful cyberoperations, in part because classifications created in physical space do not map well onto the cyber domain and in part because states appear unwilling to extend existing definitions of unlawful state acts. As a result, states victim to injurious and invasive cyberoperations currently have few non-escalatory responsive options, and the harms associated with these incidents tend to lie where they fall.

To address this growing issue, this Article draws on tort law and international law principles to construct a comprehensive regime of state accountability in cyberspace, where states are both liable for their harmful acts and responsible for their wrongful ones. It identifies international cybertorts—acts that employ, infect, or undermine the internet, a computer system, or a network and thereby cause significant transboundary harm—as a distinct kind of cyberoperation, and in doing so distinguishes and clarifies the boundaries of cybercrime and cyberwarfare. Recognizing this new category also highlights how the principle of state liability for transboundary harms (which holds states accountable for the harmful consequences of their lawful and unlawful activities) could usefully complement the existing law of state responsibility (which applies only to unlawful state actions) to create term “terrorism,” it is controversial whether a state can engage in cyberterrorism.


16 CYBERSECURITY VENTURES, CYBERSECURITY MARKET REPORT (2016), http://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/ (noting that this assessment includes “includes damage and destruction of data, stolen money, lost productivity, theft of intellectual property, theft of personal and financial data, embezzlement, fraud, post-attack disruption to the normal course of business, forensic investigation, restoration and deletion of hacked data and systems, and reputational harm”).
a comprehensive system of state accountability in cyberspace. Additionally, imposing state liability for international cybertorts preserves a bounded grey zone for state experimentation in cyberspace, minimizes the likelihood that victim states will resort to escalatory responses, and increases the chance that those harmed by cyberoperations will be compensated.

There are many reasons states could be expected to welcome these clarifications. Delineating international cybertorts creates an intermediate category between unproblematic state activity in cyberspace and cyberwarfare. Labeling a harmful cyberoperation as an international cybertort does not mean that it was necessarily unlawful; rather, it puts the perpetrator on notice that they might be liable for compensating associated injuries. Nor is this proposal an entirely revolutionary idea in international law: various treaties describe liability standards for different kinds of conduct, and states regularly set up institutions to evaluate state liability for harms and settle claims for compensation.

While these proposed concepts could be incorporated within the existing international enforcement mechanisms, states would ideally create an independent institution with the expertise and investigative resources to impartially assess state accountability in cyberspace, the flexibility to adapt to changing technologies, and the enforcement authority to deter states from engaging in inappropriate and escalatory self-help. This would not entirely eliminate legal gray zones—new technological developments, state reluctance to disclose technological capabilities, and state interest in

17 On this point, I owe a great debt to conversations with Beatrice Walton, who has since published the first piece of scholarly writing defining the concept of state liability in international law and evaluating how it applies to cyberoperations. Note, Beatrice Walton, Duties Owed: Low-Intensity Cyber Attacks and Liability for Transboundary Torts in International Law, 126 YALE L.J. 1478-88 (2017).


19 For example, many military powerhouses already voluntarily compensate victims of their actions in armed conflicts with ex gratia payments. See, e.g., Paul von Zielbauer, Confusion and Discord in U.S. Compensation to Civilian Victims of War, N.Y. TIMES (Apr. 12, 2007), http://www.nytimes.com/2007/04/12/world/americas/12iht-abuse.1.5246758.html (noting that, between 2001 and the spring of 2007, the United States paid approximately $32 million to civilians injured or the families of civilians killed in Afghanistan and Iraq); see also Rebecca Crootof, War Torts: Accountability for Autonomous Weapons, 164 U. PA. L. REV. 1347, 1393 (2016) (discussing the U.S. Foreign Claims Act and NATO’s Status of Forces Agreement (SOFA) as examples of states voluntarily committing to compensate the victims of their activities).

In other situations, compensation claims are settled after protracted legal proceedings. For example, after five years of litigation, the United States settled a claim with Iran on behalf of the victims of the downing of Iran Air Flight 655 for $61.8 million. Settlement Agreement, Aerial Incident of 3 July 1988 (Iran v. U.S.), http://www.icj-cij.org/docket/files/79/11131.pdf.
preserving some unregulated space will ensure there is plenty of grist for the academic mill—but an independent institution would improve the likelihood that the international law of cyberspace develops in a cohesive manner.

Part I reviews how the architecture of cyberspace and the structure of the modern international legal order—particularly its restrictions on self-help measures—has resulted in a lack of effective, non-escalatory deterrents to increasingly harmful but difficult-to-classify cyberoperations. Part II identifies international cybertorts as a distinct class of cyberoperations; clarifies its relationship with cyberwarfare, transnational cybercrime, data destruction, ransomware, cyber exploitation, and cyberespionage; and proposes that states be held accountable for their cybertorts under the principle of state liability for transboundary harms. Part III reviews the law of state responsibility; discusses why states are more likely to engage in certain kinds of internationally wrongful acts given the characteristics of cyberspace, and argues for minimizing resort to claims of state responsibility in light of the possibility of state liability. Part IV considers how best to develop a comprehensive accountability regime for state activity in cyberspace.

I. A Problem Without a Name

The architecture of cyberspace favors attackers, preventing states from enacting effective defensives; simultaneously, existing international law limits victim states’ recourse to effective and non-escalatory ex post deterrents. As a result, state-sponsored cyberoperations are increasingly common, costly, and invasive.

For example, in October 2012, Chinese government officials warned the New York Times that its investigation into how relatives of Wen Jiabao, China’s prime minister, had recently accumulated billions would “‘have consequences.’”\(^\text{20}\) Four months later, the paper publicized that Chinese hackers had infiltrated its computer systems.\(^\text{21}\) While they could have utterly destroyed the Times’s network infrastructure, the hackers instead appeared to be looking for information as to sources in the investigation.\(^\text{22}\) The intrusion was quarantined and eliminated only after the Times hired a private company that specialized in security breaches, set up new defenses, and replaced all compromised computers.\(^\text{23}\)

\(^{21}\) Id.
\(^{22}\) Id.
\(^{23}\) Id.
In December 2014, the Sands Casino was attacked, allegedly by Iranian hackers: computers and servers shut themselves down and hard drives were wiped. This was the first known example of a cyberoperation targeting an American business designed to destroy (rather than spy or steal)—and the attack was almost undoubtedly retaliation for comments that its CEO, Sheldon Adelson, made about nuking Tehran. The immediate costs of lost equipment and data were estimated at $40 million.

On the April 2015 premiere date of TV5Monde, a new French broadcast channel, allegedly Russian cyberoperations targeted and disrupted the “Internet-connected hardware that controlled the TV station’s operations.” The financial costs of the attack ran to €5 million in 2015, and over €3 million every year since for new protections. This, however, was cheap compared to the possibility of the entire business being destroyed. TV5Monde’s director-general recalled that, due to the risks of customer cancelations, “We were a couple of hours from having the whole station gone for good.”

Other costly and intrusive state-sponsored cyberoperations include the 2008-2012 U.S.-funded Twitter-like social media platform in Cuba, the U.S.- and Israeli-linked 2010 Stuxnet attack, the 2012 Iranian-linked attack on a Saudi Arabian oil company, the 2015 Russian-linked attack on the Ukrainian electrical grid, the 2016 U.S. Internet shut down—and, of course, the 2014 Sony hack and 2016 DNC hack.

States and scholars are looking to international law for guidance on how to lawfully respond to these harmful cyberoperations. But international law has little to say on the subject—except to limit a victim state’s lawful unilateral self-help options.

A. Modern International Law's Limitations on Self-Help

Limited state recourse to self-help measures is a feature of the modern international legal order, which prioritizes international peace over perfect enforcement. States are expected to let minor slights and violations of international law go unaddressed in order to minimize perpetuating cycles of escalatory self-help.

25 Id.
26 Id.
28 Id.
29 Id.
This was not always the case. Historically, international law was created and enforced through self-help measures. “Self-help” refers to “private actions taken by those interested in [a] controversy to prevent or resolve disputes without official assistance of a governmental official or disinterested third party.”³⁰ The legitimacy of self-help has long been recognized, particularly in environments where there is no authoritative lawmaker or law enforcer.³¹ Accordingly, the early international legal system was one in which states would often use military force to settle a wide range of disputes.³²

Allowing individual actors to unilaterally address wrongdoing has a number of benefits: it “may serve to deter such wrongdoing from occurring in the first place, reduce administrative costs, promote autonomy- or sovereignty-related values, and facilitate speedier redress.”³³ At a larger level, self-help “might serve to facilitate the maintenance of cooperative relations, mitigate feelings of alienation from the law, or generate deeper internalization of first-order legal norms.”³⁴

Self-help systems, however, are inherently unstable and prone to conflict escalation. Because self-helpers judge their own cause, “[t]here is ample reason to worry that they will misconstrue the law along the way—not just, or even primarily, on account of bad faith, but on account of motivated cognition and reliance on congenial interpretive methods or theories of law.”³⁵ Self-help regimes also disproportionately favor the powerful and foster vicious circles of attacks and counterattacks.

Given the likelihood that self-help will result in inappropriate responses and conflict escalation, legal systems often develop in part to minimize recourse to self-help. This was an animating reason for the formation of the U.N. Charter, which sharply restricts the use of violent self-help, as well as the development of the law of countermeasures, which limits state recourse to non-violent self-help measures.

1. Charter Restrictions on the Use of Force

Article 2(4) of the U.N. Charter prohibits states from unilaterally using force: “All Members shall refrain in their international relations from the

---

³¹ Id. at 843 (“Prior to the existence of legal institutions to dictate rules of behavior and state authorities to enforce them, all social relations were a form of self-help.”).
³⁴ Id.
³⁵ Id. at 50.
threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations.”

Instead of taking matters into their own hands, states are expected to pursue institutionalized means of resolving major disputes—for example, states may lawfully use force against another state after having procured an authorizing Security Council resolution—and let minor ones go unpunished.

There is one express exception to Article 2(4)’s general prohibition on unilateral state recourse to force. Article 51 provides: “Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security.” The threshold for an “armed attack” is generally understood to be higher than that required for a use of force, though the U.S. minority opinion that the two terms are essentially co-extensive has created a legal debate that has extended to assessments of the law of cyberspace.

2. Customary Limits on the Use of Countermeasures

The law of countermeasures developed in the shadow of the U.N. Charter as a means by which states victim to “below the threshold” acts could still take unilateral action to bring international law violators back into compliance without contravening Article 2(4). Countermeasures are “measures that would otherwise be contrary to the international obligations of an injured State vis-à-vis the responsible State, if they were not taken by the former in response to an internationally wrongful act by the latter in order to procure cessation and reparation.” For example, an injured state may suspend transit or trade rights with a state in violation of a treaty until it ceases

---

37 U.N. Charter art. 39.
38 Id. art. 51.
40 See, e.g., Remarks by Harold Hongju Koh, International Law in Cyberspace, USCYBERCOM Inter-Agency Legal Conference, Ft. Meade, MD (September 18, 2012).
41 TALLINN MANUAL 2.0 ON THE INTERNATIONAL LAW APPLICABLE TO CYBER OPERATIONS 333 (Rule 69) cmt. 7 (Michael N. Schmitt ed., 2017) [hereinafter TALLINN MANUAL 2.0] (acknowledging the U.S. minority view).
the wrongful act or makes appropriate reparation.\textsuperscript{43}

In keeping with the U.N. Charter’s general bent towards minimizing escalation, violent countermeasures are not permitted: “Countermeasures shall not affect . . . the obligations to refrain from the threat or use of force as embodied in the Charter of the United Nations.”\textsuperscript{44} Thus, reciprocal violent countermeasures in response to a violation of Article 2(4) would also violate Article 2(4).\textsuperscript{45}

Furthermore, even the use of non-violent countermeasures is strictly circumscribed: there are many situations in which they cannot be used at all,\textsuperscript{46} and when countermeasures are allowed, they must satisfy a number of prerequisites to be lawful. To name just a few, countermeasures must be temporary in their effects, comply with the principles of necessity and proportionality, and be designed to induce compliance with international law. Importantly, countermeasures must be taken “to procure cessation and reparation” of an internationally wrongful act—not to punish.\textsuperscript{47}

In contrast to the strict restrictions on the use of force and countermeasures, states may always employ retorsions to attempt to alter another state’s behavior. While countermeasures are acts that would be unlawful but for the fact that they are taken to restore order (like reciprocal treaty breaches), retorsions are politically unfriendly but lawful self-help measures (like discontinuing development aid, declaring a diplomat \textit{persona non grata}, or imposing unilateral sanctions).\textsuperscript{48}

\textsuperscript{43} Id. at 130.

\textsuperscript{44} Id. art. 50(a); see also id. art. 59 (“These articles are without prejudice to the Charter of the United Nations.”); id. at 327 (noting that the Articles apply only to “non-forceful countermeasures”); id. at 334 (excluding “forcible measures from the ambit of permissible countermeasures”); see also TALLINN MANUAL 2.0, supra note 41, at 125 (Rule 22) cmt. 11 (noting that the majority of experts consider “the obligation to refrain from the use of force” to be “a key limitation on an injured State when conducting countermeasures”).

\textsuperscript{45} Oona A. Hathaway, \textit{The Drawbacks and Dangers of Active Defense}, in 6TH INTERNATIONAL CONFERENCE ON CYBER CONFLICT PROCEEDINGS, NATO CCD COE PUBLICATIONS 39, 48 (2014). \textit{But see} TALLINN MANUAL 2.0, supra note 41, at 125 (Rule 22) cmts. 12-14 (noting that the minority opinion that “forcible countermeasures are appropriate in response to a wrongful use of force that itself does not qualify as an armed attack” and finding support in Judge Simma’s separate opinion in the \textit{Oil Platforms} judgment).

\textsuperscript{46} For example, a victim state may not use countermeasures in response to an internationally wrongful act that has ceased and is unlikely to be repeated, Draft Articles, supra note 42, arts. 49(2), 52(3)(a); when the internationally wrongful act has ended and the issue is pending before a third-party dispute settlement procedure, id. art. 52(3); and countermeasures cannot violate fundamental human rights, \textit{jus cogens} norms, the prohibition on belligerent reprisals, or dispute settlement procedures, id. art. 50(1).

\textsuperscript{47} Id. at 128.

In short, states may unilaterally use defensive force in response to armed attacks, states may sometimes engage in countermeasures to correct another state’s unlawful acts, and states may always employ retorsions in the attempt to alter another state’s behavior. However, these rules and enforcement mechanisms were developed in the physical world and founded on assumptions that do not translate well to cyberspace—resulting in a general lack of credible deterents to harmful state-sponsored cyberoperations.

B. The Need for Effective, Non-Escalatory Deterrents

Deterrence theory is based on the presumption that certain actions will either be unsuccessful or lead to consequential and painful responses. But states victim to harmful or invasive cyberoperations have few lawful responsive options. Attempts at deterrence-by-denial in cyberspace are of limited utility, as contemporaneous responsive measures can only do so much in an environment that favors attackers. Meanwhile, international legal restrictions on self-help measures become even more restrictive when translated to cyberspace, limiting victim states’ unilateral ex post options.

1. Practical Limits of Deterrence-by-Denial

A state or non-state entity should be expected to take basic precautions to avoid being too easy a target. The United States, for example, has emphasized the importance of “deterrence by denial”—in other words, engaging in better cybersecurity practices and beefing up defenses.

Deterrence by denial, however, offers limited protection in the cyber context. Defenders are playing an elaborate game of whack-a-mole, where a single missed attack can have devastating effects. Further, while cybersecurity good practices are important and while there are some justifications for holding states accountable for egregiously poor cybersecurity, overemphasizing a due diligence requirement risks focusing

\[\text{news/2015/jan/02/obama-imposes-sanctions-north-korea-sony-hack-the-interview.}\]

\[\text{49 The most extreme version of this is “mutually-assured destruction,” where the threat of a full-scale use of nuclear weapons by opposing sides would result in the complete annihilation of both (and, possibly, the rest of the world).}\]

\[\text{50 At present, this is not a recognized formal duty under international law, though there is growing sympathy for the concept of cyber due diligence—a requirement to not allow harm to emanate from state territory. Should a norm of cyber due diligence be recognized, cybersecurity measures might be requirements (as opposed to best practices). See infra Part II(B)(3)(c).}\]

more on the culpability of the entity that left the door unlocked than of the entity that trespassed and burglarized the building.

Lastly, overreliance on contemporaneous defense invites many of the problems associated with self-help measures. The speed of cyber will nearly always require that in-the-moment defenses be automated or autonomous. In most circumstances, a particular cyberoperation will be neutralized by a defense system long before a human being even knows it was attempted. As long as defensive measures are primarily passive and simply attempt to shield the target network or repair damage, the lack of human input is relatively unproblematic. But active defenses are a different story. Active defenses can be loosely defined as “a set of operational, technical, policy, and legal measures” that “captures a spectrum of proactive cybersecurity measures that fall between traditional passive defense and offense.” These might include both “technical interactions between a defender and an attacker” and “operations that enable defenders to collect intelligence on threat actors and indicators on the Internet, as well as other policy tools (e.g. sanctions, indictments, trade remedies) that can modify the behavior of malicious actors.” At the far end of the spectrum, some active defenses risk crossing the line into uses of force. If defenses are active and entirely automated or autonomous, it is easy to imagine an exchange of attacks and counter-attacks that quickly escalates into warfare.

Nor are states the only players. Given that the majority of cyberoperations target private sector entities, some cybersecurity experts are suggesting that industries take a more proactive approach. The more non-state entities employ active defenses without guidance from states, however, the more likely they are to respond in ways that implicate national security and foreign

and technological responsibilities on states that are or may be the target of harmful cyberoperations).


53 ACTIVE DEFENSE REPORT, supra note 15, at 1, 9 (emphasis omitted).

54 Id. at 9 (emphasis omitted).

55 This was a foundational assumption in an entire genre of science fiction novels written or movies produced during the Cold War. See, e.g., MORDECAI ROHSHALD, LEVEL 7 (1959); WAR GAMES (United Artists 1983).
relations concerns and increase the likelihood of unintended conflicts.\textsuperscript{56} For example, in response to a targeted malware attack termed Operation Aurora, Google appears to have gained unauthorized access to Taiwanese computers believed to be under the control of Chinese entities—which could be interpreted as a violation of the Computer Fraud and Abuse Act and which might have had problematic political implications.\textsuperscript{57}

2. Practical and Legal Limits of Deterrence-by-Punishment

The risks and insufficiency of deterrence by denial suggests the need for a more traditional conception of “deterrence by punishment.” In other words, in the absence of effective ex ante defenses, states need strong ex post deterrents. But deterrence strategies developed in physical space do not always translate well to the cyber domain.\textsuperscript{58}

First, there is the attribution problem. If a victim state cannot quickly and reliably identify the actual perpetrator, it will not be able to take timely and appropriate responsive actions. While eventual attribution of a cyberoperation is becoming more feasible (especially as state-sponsored cyberoperations commonly rely on previously-used architectures), it is still nearly impossible to identify the actual perpetrator immediately.\textsuperscript{59} This difficulty is compounded when a state acts through a non-state actor. Consider the attack on TV5Monde: the hackers claimed to be members of a group called the Cyber Caliphate and implied that they were linked to the Islamic State; later evidence suggests that the actual perpetrators were a

\textsuperscript{56} Accordingly, many proposed active defenses should only be taken by private sector entities working in close collaboration with government. Nuala O’Connor, Appendix I: Additional Views of Nuala O’Connor, in \textit{ACTIVE DEFENSE REPORT}, \textit{supra} note 15, at 40.

\textsuperscript{57} \textit{Id.} at 14, 40. Given the many benefits and risks associated with private sector active defense, some have proposed a middle path, with specific limitations on and incentives for appropriate private use of active defenses. \textit{See, e.g.}, WYATT HOFFMAN \& ARIEL E. LEVITE, \textit{PRIVATE SECTOR CYBER DEFENSE: CAN ACTIVE MEASURES HELP STABILIZE CYBERSPACE?} (2017). In 2017, Representative Tom Graves released a draft “Active Cyber Defense Certainty (ACDC) Act” that would permit limited private defensive measures in cyberspace. Active Cyber Defense Certainty Act – 2.0 (Discussion Draft), \textit{available at} https://tomgraves.house.gov/uploadedfiles/discussion_draft_active_cyber_defense_certainty_act_2.0_rep_tom_graves_ga-14.pdf.

\textsuperscript{58} MARTIN LIBIKI, \textit{CYBERDETERRENCE AND CYBERWAR} (2009).

group of Russian hackers known as APT 28 or “Fancy Bear” (the same group likely responsible for the DNC hack). More recently, the NATO Cooperative Defence Centre of Excellence concluded that the NotPetya malware that purported to be created by common cybercriminals “was probably launched by a state actor or a non-state actor with support or approval from a state.” In support of this claim, it argued that the operation was likely too complex to have been orchestrated by unaffiliated hackers, and that the ransomware collection method “was so poorly designed that the ransom would probably not even cover the cost of the operation.” However, at the time of the writing, it had not publicly attributed the cyberoperation to a specific state.

Even if both the cyberoperation and perpetrator are reasonably identifiable, the intended effect or message of a cyberoperation is often unclear. For example, in August 2016, a Twitter account named “theshadowbrokers” associated with state-sponsored Russian hackers, posted a link to a cache of computer codes outlining hacker tools allegedly stolen from the Equation Group, a hacker group long associated with the U.S. National Security Agency. In a series of tweets, Edward Snowden discussed why this disclosure might have far-reaching foreign policy implications:

> Circumstantial evidence and conventional wisdom indicate Russian responsibility [for the hack]. Here’s why that is significant: This leak is likely a warning that someone can prove US responsibility for any attacks that originated from

---


61 NATO CCD COE, supra note 100.

62 Id.; see also id. (quoting Laui Lindström, a NATO CCD COE Strategy Branch researcher, as stating that NotPetya is “likely . . . a declaration of power – demonstration of the acquired disruptive capability and readiness to use it’’).

63 Others have taken the analysis a step further, arguing that given that 60% of infected machines are in Ukraine and that the attack began the day before the Ukrainian Constitution Day, the attack was likely politically motivated. Lee Mathews, *The NotPetya Ransomware May Actually Be A Devasting Cyberweapon*, FORBES, Jun. 30, 2017, https://www.forbes.com/sites/leematthews/2017/06/30/the-notpetya-ransomware-may-actually-be-a-devastating-cyberweapon/#6ef1c94f39e8. While this has caused some to suggest Russia was responsible for the attacks, others are more reserved: Brian Lord, former deputy director for intelligence and cyber operations at the U.K. Government Communication Headquarters and currently the managing director for cyber and technology at PGI Cyber, noted, “There’s something about the blatanness of hitting Ukraine that doesn’t sit well with me about this being a Russian attack.” Sheera Frenkel, Mark Scott & Paul Mozur, *Myster of Motive for a Ransomware Attack: Money, Mayhem or a Message?*, N.Y. TIMES, Jun. 28, 2017, https://www.nytimes.com/2017/06/28/business/ramsonware-hackers-cybersecurity-petya-impact.html?mcubz=0.
This leak could have been, as Snowden hypothesized, a threat or form of blackmail; it could also have been a provocation, information warfare intended to discredit the NSA, propaganda for Russian or other audiences, or a response to some other action (either in cyberspace or physical space, publicly known or not). The appropriate response would differ depending on what the action was or was meant to be—but without knowing the intent of the action, it is difficult for states to take an appropriate response. Furthermore, while the disclosure of the NSA hack was apparently deliberate, it is possible to imagine a scenario where malware intended for cyberespionage malfunctioned and caused some unintended, harmful result. In an ideal world, the victim state would react differently to a mistake than to intentional destruction; in this one, a victim state might not be able to make that distinction.

In situations where a victim state knows it has been subject to an invasive cyberoperation and can reasonably identify both the state perpetrator and the purpose of the action, there are fewer lawful responsive options available in the cyber domain than in physical space. As discussed above, modern international law discourages violent state vigilantism and limits a state’s ability to engage in lawful ex post punishments to avoid the risk of conflict escalation. States are permitted to engage in unilateral self-help subject to strict legal limitations—but those limitations prevent most forms of state self-help in response to harmful cyberoperations. Should a cyberoperation constitute an armed attack, the victim state can use defensive force in response, but most cyberoperations are not sufficiently destructive to meet the armed attack threshold. States victim to “below the threshold” cyberoperations may theoretically employ countermeasures and retortions to

---


65 See supra Part I(A); see also TALLINN MANUAL 2.0, supra note 41, at 116 (Rule 20).

66 See infra text accompanying notes 99-103.
alter a suspected perpetrator’s behavior, but absent amendment or significant reinterpretation, these options have limited efficacy as deterrents.

Because self-help measures are meant to be a last resort, countermeasures are only supposed to be employed after submitting a formal request to the responsible State to remedy its internationally wrongful act (subject to certain exceptions). To avoid perpetuating cycles of escalatory self-help, a state may not employ countermeasures after an internationally wrongful act has ceased, and punitive countermeasures are prohibited. However, the speed and secrecy of cyber means that many harmful acts will have ended before they are discovered, let alone before the victim state is able to identify the responsible state and issue a request for cessation or employ a timely countermeasure.

Additionally, states use countermeasures at their own risk. Recall that countermeasures are otherwise-unlawful acts that are only permissible when used to induce another state’s compliance with its international obligations. Should a state use countermeasures inappropriately or against the wrong entity, the original victim state becomes responsible for an internationally wrongful act. As there are myriad opportunities for victim states acting in good faith to misidentify perpetrators and for state and non-state-actors to launch cyberoperations that encourage such misidentifications, states may be hesitant to employ countermeasures until they are reasonably certain of the perpetrator. Between uncertainty about what constitutes an internationally wrongful act and uncertainty about being able to make a reasonable attribution, states are likely to have delayed reactions to cyberoperations—and delayed reactions look more like prohibited punishment than permissible countermeasures.

Part of the problem is that countermeasures were never meant to be deterrents: at least theoretically, they are formally restricted to being used to restore the status quo prior to the perpetrator’s legal violation. In physical space, where many unlawful acts are public, relatively easily attributable, and take place over an extended period of time, the line between a victim state attempting to restore order and taking retaliatory action was blurred, allowing

---


68 Draft Articles, supra note 42, arts. 43(2), 52(1); id. at 135.

69 See id. arts. 49(2), 50(1), 52(3).

70 But see TALLINN MANUAL 2.0, supra note 41, at 116 (Rule 20) cmt. 16 (noting a minority view that honest and reasonable mistakes are not unlawful).

71 Id. at 128.
lawful countermeasures to also serve as functional deterrents. This line is far crisper in cyberspace, rendering most countermeasures unlawful or of little use in addressing ongoing cyberoperations or deterring future ones.

Of course, states may always employ retorsions as a deterrent or punishment. Importantly, neither countermeasures nor retorsions need mirror the actions they are intended to stop or deter—a state victim to a harmful cyberoperation could respond with economic sanctions (as the United States did with North Korea following the Sony hack) or by ousting diplomats (as the United States did to Russian officials following the DNC hack). But these unilateral retorsions have not been effective at deterring malicious cyberoperations.

3. State Paralysis

In the absence of clear rules delineating lawful and unlawful state behavior in cyberspace, victim states appear unsure of how to respond to harmful cyberoperations; even when they are reasonably certain of the perpetrator’s identity, it is not clear what responsive measures they may take as a matter of law or should take as a matter of policy.  

At present, victim states seem to be erring on the side of minimal public action, possibly to avoid setting undesirable precedent or risking uncontrolled conflict escalation. Consider the delayed U.S. reaction to the DNC hack. Although Russian involvement was suspected from the outset—and, as has been subsequently disclosed, the United States had information that the Russian government had accessed the Democratic National Committee networks as early as July 2015—the United States did not publicly attribute the hack to Russia until October 2016. After months of speculation and proposals as to what the United States might do in response, at the end of  

---

72 But see Eric Talbot Jensen & Sean Watts, A Cyber Duty of Due Diligence: Gentle Civilizer or Crude Destabilizer?, 95 TEX. L. REV. 1555, 1564 (2017) (suggesting that state’s minimalistic response to malicious cyberoperations is due more to uncertainty regarding attribution rather than uncertainty regarding the applicable law).

73 Susan Hennessey has described this situation as a “‘paralysis of too many options.’” Isabella Uria, Hacking the Election Conference, at 2 (2016), http://isp.yale.edu/sites/default/files/hacking_the_election_conference_report_11.01.16_0.pdf.


75 Sanger & Savage, supra note 6.

December 2016 the Obama Administration imposed sanctions on four individuals and five Russian entities, expelled 35 suspected Russian intelligence operatives, shut down two U.S.-based Russian compounds, and released information on Russian cyber tactics and techniques. Collectively, these actions constituted the strongest U.S. public response to a cyberoperation to date; nonetheless, they were widely derided as being insufficient to deter similar future cyberoperations. Many are concerned that the U.S. “pattern of vacillation in response to very damaging cyberoperations will not deter adversaries; it will embolden them.” Of course, the United States is not the only state victim to harmful cyberoperations; numerous other states have been subjected to similar actions (and the United States is widely recognized as a perpetrator). But this minimalist response is common—fostering a friendly environment for unchecked state-sponsored cyberoperations.

What is needed is a new ex post deterrent, crafted in light of states’ new involvement in the attacks” with the hope of eventual U.N. condemnations and economic sanctions; “undermine the Russian government’s reliance on a wide variety of cyber-tools to censor the web within its own country,” “expose the overseas banking accounts and financial resources of high-level Russian government officials,” “punish Russian hackers by knocking them off-line or even damaging their hardware,” or turn to allies for help).


See, e.g., Rebecca Crootof, The DNC Hack Demonstrates the Need for Cyber-Specific Deterrents, LAWFARE (Jan. 9, 2017, 8:00 AM), https://www.lawfareblog.com/dnc-hack-demonstrates-need-cyber-specific-deterrents (providing examples).


As Sean Watts has observed, “[i]n addition to being highly feasible and often inexpensive, low-intensity cyber operations offer attractive prospects for anonymity, appear to frustrate attack correlation by targets, and may also reduce the likelihood of provoking severe retaliation. In short, low-intensity cyber operations offer States appealing opportunities to degrade adversaries while avoiding the likely strategic and legal costs of massively destructive cyber attacks.” Sean Watts, Low Intensity Cyber Operations and the Principle of Non-Intervention, in CYBER WAR: LAW AND ETHICS FOR VIRTUAL CONFLICTS 249, 250 (Jens David Ohlin et al. eds., 2015).
cyber-enabled capabilities and designed to address the harms caused by
cyberoperations like the Sony and DNC hacks.\textsuperscript{83} Accordingly, this Article
proposes recognizing international cybertorts as a natural complement to
cyberwarfare and transnational cybercrime and distinguishing it from
inherently internationally wrongful acts. Once these different types of
cyberoperations are delineated, it is possible to construct tailored
accountability regimes that permit state experimentation in cyberspace,
provide victim states with non-escalatory deterrent responses, and
compensate victims.

II. STATE LIABILITY FOR INTERNATIONAL CYBERTORTS

This Article defines an international cybertort as an act that employs,
infests, or undermines the internet, a computer system, or a network and
thereby causes significant transboundary harm. Not only is it conceptually
useful to differentiate international cybertorts from transnational cybercrime,
cyberwarfare, and other kinds of cyberoperations, this new term implies an
alternative accountability mechanism and deterrent: states could be required
to compensate victims of their international cybertorts under the principle of
state liability for transboundary harms.

A. A Distinct Kind of Harmful Cyberoperation

As noted above, the North Korean “Guardians of Peace” hacker group
raided Sony Entertainment Pictures servers and publicized extensive
confidential data.\textsuperscript{84} Based on an FBI analysis of the associated software,
techniques, and network sources, the United States took the then-
unprecedented move of publicly attributing the Guardian’s cyberoperations
to the state of North Korea.\textsuperscript{85} Officially, the United States retorted by
imposing new unilateral sanctions\textsuperscript{86}; many suspect that the United States was
also responsible for extensive North Korean web outages in early December
2014.\textsuperscript{87} As of February 2015, Sony estimated that investigation and
remediation costs had reached $15 million—ultimately, the total direct and
indirect costs of the hack will likely be far greater, with some estimating that
costs might include $70 million in direct damages and another $130 million

\begin{itemize}
\item \textsuperscript{83} Crootof, \textit{supra} note 79.
\item \textsuperscript{84} Robb, \textit{supra} note 1.
\item \textsuperscript{85} Office of the Press Center, \textit{supra} note 2.
\item \textsuperscript{86} Roberts, \textit{supra} note 48.
\item \textsuperscript{87} Strohm, \textit{supra} note 11.
\end{itemize}
in indirect damage (such as leaked trade secrets and lost revenue).  

To the extent it was conducted by or sponsored by a state, the Sony hack was not—or was not only—a transnational cybercrime.  

A cybercrime occurs when a computer or program is used as the means to commit an illegal act.  

Domestic cybercrimes are regulated internally; cross-border cybercrimes are investigated and prosecuted like other kinds of transnational crime. A paradigmatic example of transnational cybercrime occurred in August 2015, when Wall Street traders partnered with Ukrainian hackers to gain access to unpublished company press releases, allowing them to make trades that “reaped more than $100 million in illegal proceeds.”  

Significantly, only individuals are subject to criminal liability for cybercrimes—states cannot be held criminally liable, even for state-sponsored cybercrimes. Thus, while the Guardians of the Peace or identified state actors or agents could (theoretically) be criminally prosecuted for the Sony hack, the state of North Korea cannot.  

Nor was the Sony hack an act of cyberwarfare. Although a cyberoperation might be intended to undermine a state’s national security, be politically coercive, or cause extensive economic harm, such an action only constitutes “cyberwarfare” if it is sufficiently destructive in physical space to meet the “armed attack” threshold legitimizing defensive military action or if it occurs in the context of an ongoing armed conflict.  

While no cyberoperation has yet initiated an armed conflict, they have been used in response to traditional provocations or in conjunction with more conventional attacks. For example, the United States infiltrated the Iraqi Defense Ministry email system to inform Iraqi officers how they could peacefully surrender shortly before its 2003 invasion; the Israeli Air Force used a cyberoperation to compromise the Syrian air-defense system during

---

88 Brinded, supra note 3.


90 See id. at 833-34; see also Convention on Cybercrime, supra note 9, pmbl.


92 U.N. Charter art. 51; Hathaway et al., supra note 67, at 839-40 (discussing cyberwarfare).

its 2007 air strike against a nuclear facility; and, in the summer of 2008, Georgia’s internet access was shut down while Russian forces invaded South Ossetia. More recently, the United States has publicly announced that it is using cyberoperations in its campaign against the Islamic State, both to interfere with its communications strategy and to alter data in its systems. “We are dropping cyberbombs,” said Deputy Secretary of Defense Robert Work, “We have never done that before.” Jus in bello (the law governing the conduct of hostilities) regulates cyberoperations that occur in the context of an armed conflict—but the Sony hack occurred during peacetime.

There is general agreement that jus ad bellum (the law governing the commencement of hostilities) regulates responses to cyberoperations that satisfy the armed attack threshold requirement, and that a cyberoperation can meet that standard if its effects are equivalent to those of kinetic armed attack. However, the vast majority of harmful cyberoperations cause non-physical damage that simply doesn’t register on the cyberwarfare spectrum. Notwithstanding perennial academic interest in the question, at present only one cyberoperation—the 2010 Stuxnet attack—has arguably had sufficiently

---

97 Id.
98 TALLINN MANUAL 2.0, supra note 41, at 375 (Rule 80) cmt. 10; Hathaway et al., supra note 67, at 839-56.
99 See, e.g., TALLINN MANUAL 2.0, supra note 41, at 339 (Rule 71); see also id. at 328-29 (Chapter 14) cmt. 3 (noting that the Tallinn Manual 2.0 applies the lex lata norms in the cyber context, but that these are subject to change based on state practice).
100 Id. at 333-37 (Rule 69) cmts. 8-11; Hathaway et al., supra note 67, at 836-37. Notably, in 2014 NATO declared that it would consider a cyberoperation that rose to the level of an armed attack on one of its member states to trigger Article 5, the collective defense clause. NATO’s Secretary General reiterated this statement in response to the June 2017 NotPetya malware attacks. NATO CCD COE, NotPetya and WannaCry Call for a Joint Response From International Community, Jun. 30, 2017, https://ccdcoe.org/notpetya-and-wannacry-call-joint-response-international-community.html.
destructive effects to meet the armed attack threshold. The Stuxnet attack, which destroyed 1000 Iranian centrifuges used to enrich uranium, was the first time computer malware was recognized as capable of specifically targeting and destroying industrial systems. Despite the military nature of the target and the extent of the damage, experts continue to disagree as to whether even Stuxnet qualified as an armed attack. With no military nexus and no loss of life, the Sony hack doesn’t come close to meeting the armed attack threshold.

So what was the Sony hack? Prominent writers sometimes refer to it as a “below-the-threshold” cyberoperation (referring to the armed attack threshold) or a “low-intensity cyber operation,” defined as “actions taken short of destructive or violent attacks.” Such characterizations imply that these cyberoperations exist in the negative space surrounding the law of armed conflict, but do not recognize them as a distinct kind of cyberincident with a different primary harm. In recognition of the economic damage associated with cyberoperations like the Sony hack, this Article suggests switching legal frameworks entirely—to tort law.

B. International Cybertorts

The Sony hack is best understood as an international cybertort. As opposed to the law of war, which does not consider the impact of extensive economic harms, or criminal law, which is designed to hold individuals accountable for their morally blameworthy wrongs, tort law is a regulatory tool designed to allocate liability for myriad kinds of intended and unintended injuries. Recognizing international cybertorts as a distinct category allows for a more accurate assessment of the harms associated with different cyberoperations, and by extension, a more considered discussion of how to

---

101 To date, only a handful of cyberoperations are known to have caused physical damage. These include the 2010 Stuxnet malware that destroyed Iranian centrifuges; the 2012 attack on the Saudi Arabian oil company, Aramco, which destroyed 30,000 computers; the 2014 hack and destruction of a German steel mill; the destruction of 3,000 computers and 800 servers during the 2014 Sony hack; the 2015 attacks that shut down the Ukrainian power grid; and possibly the 2008 explosion of a Turkish oil pipeline.


103 TALLINN MANUAL 2.0, supra note 41, at 342 (Rule 71) cmt. 10.


105 Watts, supra note 82.

106 The Sony hack’s relationship with cyberespionage, sovereignty violations, interference, and intervention is discussed below.
construct appropriate state accountability regimes.\textsuperscript{107}

Like the definition of transnational cybercrime, the international cybertort definition is means-based and encompasses a broad range of harmful activities.\textsuperscript{108} Rather than focusing on the intent of the actor—always a thorny question in contexts where states are legal actors\textsuperscript{109}—this definition focuses instead of the effects of the action. In other words, it is more concerned with the injury sustained by the victim is of more import than the intention of the cybertortfeasor, so the definition therefore encompasses both intentional and unintentional harms.

Importantly, the original state conduct need not itself be unlawful—rather, it is the resulting harm that raises the possibility of tort liability. This is often the case in domestic law: driving an automobile or using dynamite are lawful, albeit regulated, activities—when their use causes harm, however, the user may be liable in tort. As a result, this definition covers far more activities than would be addressed under the law of state responsibility. Furthermore, this implies that states cannot always use countermeasures in response to international cybertorts, as the original injurious activity might not be unlawful.

1. Relationship with Cybercrime and Cyberwarfare

In many domestic legal systems, the same action may sometimes be both a crime and a tort: similarly, a transitional cybertort may sometimes also be a transnational cybercrime or, if the harm is sufficiently destructive, cyberwarfare. For example, the Stuxnet attack destroyed roughly one thousand Iranian centrifuges. Assuming Iran was interested in pressing the issue, it has a credible case that this harm constituted an international cybertort. If Stuxnet was the work of non-state actors, it might also be a transnational cybercrime; if it could be sufficiently attributed to a state, it could arguably be considered cyberwarfare.

Figure 1 summarizes the similarities and distinctions between these different kinds of cyberoperations; Figure 2 illustrates the relationships between international cybertorts, transnational cybercrimes, and cyberwarfare, especially the fact that these categories are not mutually exclusive. It locates the Sony hack on the intersection of cybercrime and

\textsuperscript{107} Domestic cybertorts, in contrast, are grounded in and therefore limited by domestic tort law: they often encompass any situation where the internet or computer systems are used negligently or intentionally to commit a civil wrong.

\textsuperscript{108} In contrast, the definition of cyberwarfare is objective-based. Cf. Hathaway et al., supra note 67, at 826-28.

international cybertorts, to encompass both the criminal liability of the individual hackers who engaged in the attack and the liability of the state that allegedly sponsored the attack.

Fig. 1: Types of Cyberoperations

<table>
<thead>
<tr>
<th></th>
<th>International Cybertort</th>
<th>Transnational Cybercrime</th>
<th>Cyberwarfare</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distinguishing Characteristics</strong></td>
<td>The original cyberoperation need not be unlawful, but the act must cause significant transboundary harm.</td>
<td>Must be a violation of criminal law and committed by means of a computer system.</td>
<td>Must have a political or national security purpose and either (1) be sufficiently destructive to satisfy the “armed attack” threshold; or (2) occur in the context of an armed conflict.</td>
</tr>
<tr>
<td><strong>Governing Legal Regime</strong></td>
<td>State Liability for Transboundary Harm(^\text{10})</td>
<td>Domestic and International Criminal Law</td>
<td>(1) <em>Jus ad Bellum</em>; and (2) <em>Jus in Bello</em> (International Humanitarian Law)</td>
</tr>
<tr>
<td><strong>Paradigmatic Example</strong></td>
<td>Sony Hack</td>
<td>Ukrainian Hacker Insider Trading Ring</td>
<td>(1) possibly Stuxnet; (2) U.S. cyberoperations against the Islamic State</td>
</tr>
</tbody>
</table>

\(^{10}\) See *infra* Part II.C.
2. Relationship with Data Destruction and Ransomware

Recognizing international cybertorts as a distinct category suggests a solution for a question currently vexing cyberwarfare scholars: how to classify cyberoperations that affect access to data, either by destroying it or by holding it hostage. These attacks have myriad costly, chaotic, and even deadly effects, limited only by one’s imagination. Academic medical research centers or pharmaceutical companies could have years of trials wiped out; registered voters could be removed from the rolls; a lifetime’s worth of credit-building could disappear overnight; selective deletions of flight plans could lead to in-air collisions; the entire stock market could be thrown into chaos; information on life-threatening allergies could be removed from medical records.

Thus far, data-destruction cyberincidents have not caused any physical effects on these scales, but they have demonstrated the potential scope of the risk. In a 2011, for example, “half of the 500-plus servers belonging to [South Korea’s Nonghyup Bank] were crippled [by a data-destruction attack], including servers controlling ATMs, credit card access, and online

---

111 This Figure is intended to clarify the relationship among the different labels for different kinds of cyberoperations; it is not meant to identify which legal regime will govern in any particular situation.
banking.”\textsuperscript{112} This affected approximately 30 million customers, leading to “more than 30,000 customer complaints and 1,000 compensation claims.”\textsuperscript{113} More recently, the June 2017 NotPetya malware—so named because it presents as Petya ransomware, though it appears only capable of rendering data completely inaccessible—spread across the globe.\textsuperscript{114} During its attack, A.T.M.s stopped working, banks were forced to close, hospitals canceled operations, and the radiation monitoring system at Ukraine’s Chernobyl Nuclear Power Plant went offline.\textsuperscript{115}

Ransomware attacks can be just as devastating. Ransomware infiltrates network systems, encrypts vital files and data, and demand payment in return for the key to unlocking the information. The data in question is never destroyed, but it is rendered effectively nonexistent. Hospitals’ patient data is often targeted, as are companies’ files with their customers’ credit card data.\textsuperscript{116} A cybersecurity expert estimated that cybercriminals made over $1 billion in 2016 alone from ransomware.\textsuperscript{117} In early 2017, in the largest ransomware assault to date,\textsuperscript{118} a variant of the WannaCry ransomware “crippled 200,000 computers in more than 150 countries.”\textsuperscript{119} It “forc[ed] Britain’s public health system to send patients away, [froze] computers at Russia’s Interior Ministry and [wrought] havoc on tens of thousands of computers elsewhere.”\textsuperscript{120}

The concept of international cybertorts provides a means of holding states


\textsuperscript{113} Id.


\textsuperscript{115} Id.


\textsuperscript{119} Russell Goldman, What We Know and Don’t Know About the International Cyberattack, N.Y. TIMES, May 12, 2017, at A9.

\textsuperscript{120} Perlroth & Sanger, supra note 118.
accountable for data destruction and ransomware. A cyberoperation that compromises hospital records, banking data, or trade secrets is not an act of war, but it also cannot be tolerated.

3. Relationship with Cyber Exploitation and Cyberespionage

Cyber exploitation is the act of gathering confidential information kept on or transiting through a computer system or network, usually secretly and often for commercial purposes; cyberespionage is cyber exploitation conducted for political or military purposes. An act may be simultaneously

---

121 To the extent international cybertorts might apply to actions taken in the context of an armed conflict, they also might help address a long-standing debate about whether or not data is a lawful target. International humanitarian law prohibits attacks on civilian objects, and the Tallinn Manual 2.0 extends this prohibition: “Civilian objects shall not be made the object of cyber attacks.” TALLINN MANUAL 2.0, supra note 41, at 434 (Rule 99) (“Civilian objects shall not be made the object of cyber attacks. Cyber infrastructure may only be made the object of attack if it qualifies as a military objective.”). However, the majority of the Tallinn Manual 2.0 experts determined that “the law of armed conflict notion of object should not be interpreted as including data.” Id. at 437 (Rule 99) cmt.6. They reasoned that “[d]ata is intangible and therefore neither falls within the ‘ordinary meaning’ of the term object, nor comports with the explanation of it offered in the ICRC Additional Protocols Commentary.” Id. In other words, data may sometimes be a lawful target. A minority disagreed, contending that “for the purposes of targeting, certain data should be regarded as an object.” Id. at 437 (Rule 99) cmt. 7. Were that not the case, they argued, the deletion of even “essential civilian data sets such as social security data, tax records, and bank accounts would potentially escape the regulatory reach of the law of armed conflict.” Id. According to the minority, this would “run[] counter to principle (reflected in Article 48 of Additional protocol I) that the civilian population enjoys general protection from the effects of hostilities.” Id. While I agree with the minority, introducing the concept of international cybertorts helps alleviate the injuries that would likely accompany the majority’s approach: regardless of whether or not data is a lawful target, states might be held liable for its destruction.

I have argued elsewhere that states should be held strictly liable for the actions taken by their autonomous weapon systems. See Crotof, War Torts, supra note 19. Some aspects of this argument might be extrapolated to all accidents that occur in the context of an armed conflict.

122 There is no widely-accepted definition of either of these terms. My above definitions are intended to convey that actions constituting cyber exploitation may be directed against state or non-state actors by state or non-state actors; consist of gathering information on a computer system or network, which will usually require an electronic tool; are not intended to cause death, injury, destruction, or damage; are usually (though not necessarily) conducted secretly; often involves gathering economic information for commercial purposes; and that cyberespionage is a subset of cyber exploitations conducted to further state political and military aims. Other definitions focus on different aspects: sometimes on the parties involved, sometimes on the means of conducting the attack, and sometimes on the objectives of the attack. Cf. Seymour M. Hersh, The Online Threat: Should We Be Worried About a
cyber exploitation or cyberespionage and an international cybertorts, with the distinction turning on whether the act is discovered and publicized (by either the attacker or victim).

There are numerous examples of cyber exploitation, some of which might also constitute cyberespionage. In 2014, hackers believed to be working for the Russian government breached the White House’s unclassified network, gaining access to President Obama’s schedule and emails revealing personnel moves and policy debates.\textsuperscript{123} In March 2015, what was likely the same group hacked into the Department of State’s unclassified email network, prompting a complete system shutdown.\textsuperscript{124} Chinese hackers broke into the U.S. Office of Personnel Management (OPM) database and made off with the personal information on 22 million individuals who had undergone federal security screenings, including 4.2 million federal employees.\textsuperscript{125} These same hackers are likely also responsible for the February 2015 hack of Anthem, Inc., which compromised the personal information of roughly 80 million current and former customers, and the July and August 2015 hacks of United Airlines and American Airlines.\textsuperscript{126}

The 2016 DNC hack was a more complicated combination of cyberespionage and information warfare. In June 2016, the Democratic National Committee hired Crowdstrike to investigate a detected hack. Crowdstrike determined that the DNC had been subject to two distinct hacks, conducted by groups nicknamed Cozy Bear and Fancy Bear and associated with two organizations in the Russian government.\textsuperscript{127} Over 19,000 emails obtained through these hacks were then released by WikiLeaks.\textsuperscript{128}

\begin{itemize}
\item \textsuperscript{125} Szoldra, \textit{supra} note 123. The Chinese government has asserted that the OPM hack was individual criminal activity rather than a state-sponsored cyberoperation. Michael Forsythe & David E. Sanger, \textit{China Calls Hacking of U.S. Workers’ Data A Crime, Not a State Act}, N.Y. TIMES, Dec. 2, 2015, at A8.
\item \textsuperscript{126} Walters, \textit{supra} note 25.
\item \textsuperscript{127} Hamburger & Tumulty, \textit{supra} note 4.
\item \textsuperscript{128} Ackerman & Thielman, \textit{supra} note 4.
\end{itemize}
If the costs associated with these cyber exploitation and cyberespionage operations are sufficiently significant, they could be considered international cybertorts. And there is evidence that the economic harms are immense. With the exception of the DNC hack, these cyberoperations were likely intended to be covert—but once discovered, the hacked entities had to spend huge sums to expunge hackers, compensate customers, and rebuild reputations. Costs differ by industry: the average hack of hospital records can range from a few hundred thousand to three million dollars, while the total costs of a retailer data breach might range from two to six million dollars. Monetary costs for hacks of government entities are harder to calculate—but there are estimates that the OPM hack alone is costing taxpayers $350 million. Given that there are roughly 35,000 known computer security penetration incidents per day, the collective costs are staggering. A 2014 study estimated that cyberoperations cost businesses between $375 and $575 billion annually; a 2016 analysis suggests that costs will reach $6 trillion annually by 2021.

Distinguishing international cybertorts from transnational cybercrimes and cyberwarfare promotes a more precise understanding of the distinct kinds of harms associated with these different cyberoperations. As discussed in the next section, this new terminology also highlights the applicability of an alternative means of holding states accountable for their harmful cyberoperations, and thereby creates a new, less-escalatory responsive option for victim states.

C. State Liability

The principle of state liability for transboundary harms holds states accountable for the “injurious consequences that arise out of activities within their jurisdiction or control and that affect other States or nationals of other

---

130 Id.
132 Donlon, supra note 129.
134 CYBERSECURITY VENTURES, supra note 16.
Conceptually, this principle—that states should be held accountable for the harm they cause—already undergirds the law of state responsibility, which requires states to make appropriate reparations for their internationally wrongful acts. Consequently, state liability and state responsibility are often conflated. However, the law of state responsibility does not address harmful state operations that are not also wrongful or attributable. Given the differing harms they are meant to correct, their approaches and consequences differ: state liability is primarily concerned with ensuring compensation for injuries, while state responsibility aims to restore the status quo prior to an internationally wrongful act through a broader range of restitutive means. State liability should therefore be recognized as an independent principle, equally applicable in situations where there is no act or activity that violates a state’s international obligations.

Once distinguished from state responsibility and its requirement of an internationally wrongful act, the principle of state liability for transboundary harms creates a new responsive options for states experiencing the harmful transboundary consequences of another state’s activities.

---


136 See infra Part III(3)(A).

137 Indeed, some have suggested that the movement to distinguish the doctrine of state liability for transboundary harms is “‘fundamentally misconceived.’” Alan E. Boyle, *State Responsibility and International Liability for Injurious Consequences of Acts Not Prohibited by International Law: A Necessary Distinction?*, 39 INT’L & COMP. L.Q. 1, 13 (1990) (quoting Ian Brownlie).


140 Boyle, *supra* note 141, at 3; see also infra Part III.C.3.

141 Boyle, *supra* note 141, at 3 (“What distinguishes international liability from other forms of responsibility is that it does not presuppose wrongful conduct or breach of any obligation.”).
1. State Liability for Transboundary Harms

The intuitive idea that one is liable for caused harm is well-established in international law. It is articulated in the Roman maxim *sic utere tuo ut alienum non laedas tuo*\(^{142}\) and in Grotius’s statement that from any “Fault or Trespass there arises an Obligation by the Law of Nature to make Reparation for the Damage, if any be done.”\(^{143}\) Today, the concept of liability for caused harms is common to many domestic legal systems\(^{144}\) and reiterated in case law, treaties, and International Law Commission (ILC) reports and draft articles.

For nearly seventy years, various kinds of state liability for transboundary harms have been recognized in international jurisprudence.\(^{145}\) In the 1941 *Trail Smelter* arbitration, in which the United States claimed damages resulting from a Canadian smelter’s actions,\(^{146}\) the tribunal proclaimed that “no State has the right to use or permit the use of its territory in such a manner as to cause injury by fumes on or in the territory of another or the properties therein.”\(^{147}\) Shortly thereafter, the Permanent Court of International Justice echoed the *Trail Smelter* language in its *Corfu Channel* decision. After holding that Albania was under an obligation to warn other states that the Corfu Channel—a normally safe strait often used for international navigation—was mined, the Court stated that this obligation sprang from general international law:

The obligations incumbent upon the Albanian authorities consisted of notifying, for the benefit of shipping in general, the existence of a minefield in Albanian territorial waters and in warning the approaching British warships of the imminent danger to which the minefield exposed them. Such obligations are based, not on the Hague Convention of 1907, No. VIII, which is applicable in time of war, but on certain general and well-recognized principles, namely: *elementary considerations of humanity, even more exacting in peace*


\(^{143}\) HUGO GROTIIUS, 2 THE RIGHTS OF WAR AND PEACE 884, ¶ I (Richard Tuck ed., Jean Barbeyrar trans., 2005) (1625); see also Walton, *supra* note 17, at 1480 (discussing how, in Grotius’s time, the failure to provide compensation could itself be a just cause for war).

\(^{144}\) See ILC Survey, *supra* note 138.

\(^{145}\) See Walton, *supra* note 17, at 1478-84 (discussing relevant case law).

\(^{146}\) Sucharitkul, *supra* note 135, at 835.

\(^{147}\) *Trail Smelter* (United States v. Canada), Award, 1941, 3 R.I.A.A. 1905, 1965 (Mar. 11).
than in war; the principle of the freedom of maritime commination; and every State’s obligation not to allow knowingly its territory to be used for acts contrary to the rights of other States.\textsuperscript{148}

States have regularly concluded treaties creating or clarifying liability for specific acts or omissions.\textsuperscript{149} While passed over many decades, these treaties generally focus on a few, specific kinds of harms—injuries associated with nuclear accidents, oil spills, and other kinds of hazardous materials—or harms which endanger the use of shared spaces—international watercourses, transboundary waters, and outer space. The strongest statement of state liability is found in the 1972 Convention on International Liability for Damage Caused by Space Objects.\textsuperscript{150} It provides that “[a] launching State shall be absolutely liable to pay compensation for damage caused by its space object on the surface of the earth or to aircraft flight”\textsuperscript{151} and includes more complicated standards—including joint and severable liability and a

\textsuperscript{148} Corfu Channel (U.K. v. Alb), Judgement, 1949 I.C.J. 4, 22 (Apr. 9) (emphasis added).


\textsuperscript{151} Id. art. II. This absolute liability is subject to a defense that the damage resulted “wholly or partially from gross negligence or from an act or omission done with intent to cause damage on the part of a claimant State or of natural or juridical persons it represents.” \textit{Id.} art. VI.
contributory negligence defense—for non-Earth damage.\textsuperscript{152} Similarly, the concept of state liability for international cybertorts would help structure the use of a relatively new shared domain: cyberspace.

In 1978, the U.N. General Assembly charged the ILC “to commence work on the topics of international liability for injurious consequences arising out of acts not prohibited by international law and jurisdictional immunities of States and their property.”\textsuperscript{153} This project, originally an attempt “to conceptualize and circumscribe” state liability broadly,\textsuperscript{154} was later limited to environmental matters because of a lack of consistent state practice in other areas.\textsuperscript{155} The ILC produced two circumscribed documents: the 2001 draft articles on the duty to prevent transboundary harm from hazardous activities and the 2006 draft principles regarding liability for the injurious consequences of such actions.\textsuperscript{156} One particularly interesting conclusion of the ILC process was a clarification of the relationship between the principle of state liability and the law of state responsibility. Namely, that when a state engages in an act that is not inherently unlawful but nonetheless causes harm—say, engaging in cyberespionage—the failure to provide compensation for the harm might itself constitute an internationally wrongful act, triggering the applicability of the law of state responsibility and its broader remedial measures.

Today, the principle of state liability is primarily discussed in terms of international environmental law, largely because of the paucity of situations outside of the environmental context where state action causes significant transboundary harm. But while the doctrine of state liability for transboundary harms has been most developed in environmental law, it is

\begin{footnotesize}
\begin{enumerate}
\item Id. arts. III, IV, V, XXII.
\item ILC Report, supra note 138; ILC Survey, supra note 138, ¶ 53.
\item Sucharitkul, supra note 135, at 829.
\item Critics had two main complaints: first, they argued that there was no conceptual need to distinguish state liability from state responsibility; second, they claimed that the distinction would be of little use in developing international environmental law—and might even undermine other nascent environmental legal protections. Boyle, supra note 141, at 1.
\item Int’l L. Comm’n, Draft Articles on Prevention of Transboundary Harm from Hazardous Activities (2001); ILC Draft Principles on Allocation of Loss, supra note 170. The former concerned states’ obligations to avoid and minimize transboundary harm. The latter focused on operator liability, primarily because state liability was considered to be the exception in the treaty law. Apart from the Outer Space Treaty, most treaty regimes assigned liability for transboundary harms to non-state entities.
\end{enumerate}
\end{footnotesize}
hardly conceptually limited to environmental harms. In the *Corfu Channel* case, for example, the ICJ was not overly concerned with the environmental impact of the mines. Instead, the case was primarily about the use of military weapons in peacetime and the creation of a hazard in an otherwise safe passage used by other states. The advent of increasingly harmful state-sponsored cyberoperations and lack of effective deterrents highlights a need for a new regulatory regime—and the principle of state liability for transboundary harms provides a useful framework for thinking about state accountability in this context.

2. Benefits of State Liability for International Cybertorts

There are a number of benefits that would attend applying the principle of state liability to cyberoperations that cause significant transboundary harm. First, by allowing states to appropriately name a certain kind of harmful cyberoperation, it solves the “cybervandalism” problem: Imposing tort liability on states for the harmful consequences of their activities simultaneously permits state experimentation in cyberspace and imposes costs for harms associated with that experimentation. Second, it increases the likelihood that victims of international cybertorts will be compensated for their injuries. Finally, it suggests a new means of managing cyberespionage.

a. Creates a Non-Escalatory Responsive Option

President Obama described the Sony hack as “cybervandalism”—and was slammed domestically by Republicans for a weak characterization. But what was he to call it? Calling it cyberwarfare, as some would have preferred, would have stretched the definition of cyberwarfare beyond recognition—and would set a precedent allowing other states to term similar U.S. cyberoperations cyberwarfare.

Still, cybervandalism clearly isn’t the right term. It usually describes the annoying but relatively innocuous practice of altering online content, like website defacement. For example, in 2010 photos of Spanish Prime Minister Jose Luis Rodriguez Zapatero on a European Union website were replaced

---

157 Walton, supra note 17, at 1480.
158 Id. at 1483-84.
159 Id. at 1460.
with a close-up shot of Mr. Bean.\textsuperscript{161} Cybervandalism hardly seems to encompass the physical damage, economic costs, and reputational harms associated with the Sony hack. The language of tort law—and the term “international cybertorts”—does.

Perhaps President Obama used the term cybervandalism to avoid raising the question of the lawfulness of U.S. cyberoperations, or perhaps he used it because there was no accurate term that allowed him to denounce the Sony hack without implying the appropriateness of a military response to the hack of a civilian business.\textsuperscript{162} The concept of international cybertorts walks this thin line. Because the activity underlying an international cybertort is not necessarily unlawful, the term allows states suffering from an act’s consequences to claim compensation without prejudging its lawfulness. It simultaneously creates a new, relatively non-escalatory responsive option for a victim state: putting the alleged cyber tortfeasor on notice of potential liability.

b. Encourages Victim Compensation

A primary aim of the principle of state liability for transboundary harms is the compensation of both private and state victims. Different treaty regimes have developed different means by which those harmed by the transboundary effects of another state’s activities may bring claims for compensation: sometimes private entities bring claims against private entities, sometimes private entities bring claims against states, sometimes states bring claims against private entities, and sometimes states bring claims against states. These various approaches have corresponding benefits and drawbacks, and a careful analysis of the architecture of cyberspace and what practices have been effective in similar existing regimes is needed to determine what compensation structure will be preferable in the cyber context.

With regard to the Sony hack, the U.S. National Security Council Spokesman has already suggested that North Korea is liable for the millions in losses Sony suffered.\textsuperscript{163} However, North Korea is unlikely to pay such


\textsuperscript{162} Fung, supra note 160 (“What’s really going on here is a battle to determine whether, in fact, the infiltration of corporate networks, exposure of business information and censorship of U.S. film studios is a legitimate military activity.”).

\textsuperscript{163} Julie Makinen, North Korea Decrees U.S. Allegations on Sony hack; U.S. Turns to China, L.A. TIMES (May 5, 2016), http://www.latimes.com/world/asia/la-fg-north-korea-proposes-joint-investigation-into-sony-hack-20141220-story.html, (“[W]e are confident the North Korean government is responsible for this destructive attack . . . . If [they want] to help, they can admit their culpability and compensate Sony for the damages this attack
compensation.\footnote{Id.} This is not a damning indictment of the concept of international cybertorts, however, nor a unique phenomenon; domestic tort law plaintiffs are often unable to recoup their losses from judgement-proof defendants. While it may be sometimes impossible to compensate the victims of international cybertorts—either because there is no procedure in place for doing so, or because the perpetrator is already a rogue state and relatively immune to the threat of outcasting\footnote{See Oona Hathaway & Scott J. Shapiro, \textit{Outcasting: Enforcement in Domestic and International Law}, 121 YALE L.J. 253, 340, 341-42 (2011) (discussing how powerful and isolated states may care less about reputational costs).}—identifying international cybertorts still allows victim states to name the action and shame the liable state for not providing appropriate compensation.

c. Creates a New Means for Managing Cyberespionage

Thus far, governments have had few means of deterring cyber exploitation and cyberespionage. There is little law on the subject: the United States, one of the loudest critics of cyber-enabled industrial espionage, nonetheless maintains that “remote cyber operations involving computers or other networked devices located on another State’s territory do not constitute a per se violation of international law. . . . This is perhaps most clear where such activities in another State’s territory have no effects or de minimis effects.”\footnote{Remarks by Brian J. Egan, International Law and Stability in Cyberspace, Berkeley Law School, California (Nov. 10, 2016), available at https://www.state.gov/s/l/releases/remarks/264303.htm.}

While states are not held formally accountable for their spies’ actions in the physical world, individual spies could be apprehended, prosecuted, and punished. These physical threats helped minimize an otherwise unregulated activity. In cyberspace, spies can access far more information and data with far less personal risk, upsetting the imperfect but at least established standing equilibrium.

Even when a state identifies and reasonably attributes cyberespionage, it has few effective and non-escalatory options. Should it issue criminal indictments of suspected responsible individuals, as the United States did with members of China’s People’s Liberation Army?\footnote{Mark Landler & David E. Sanger, \textit{Hacking Charges Threaten Further Damage to Chinese-American Relations}, N.Y. TIMES, May 21, 2014, at A14.} Or should it attempt to embarrass the allegedly responsible state, by naming and shaming? Or by some covert response? All of these approaches have significant drawbacks:
criminal indictments will usually be unenforceable; naming and shaming carries little weight with regard to espionage, given that all states are likely engaged in similar activities; responsive offensive covert countermeasures are likely themselves unlawful due to the notice requirement\textsuperscript{168} and their invisibility contributes to the perception that cyberspace is a lawless zone while simultaneously risking conflict escalation.

But what if states could be held liable for the economic harm associated with discovered or publicized cyberespionage? Again, the actions undergirding an international cybertort need not be unlawful—it is the injuries associated with the activity that suggests the possibility of tortious action. The harms associated with these discovered or publicized cyberoperations are usually significant, suggesting that these acts could be identified as international cybertorts, triggering state liability and a duty to compensate. While hardly a magic bullet for the asymmetry of cyberespionage, recognizing international cybertorts as implicating state liability presents victim states with a new means of responding to it.

Clearly, there are a number of benefits that would attend expanding the principle of state liability to cyberoperations that cause significant transboundary harm. Before it can be employed, however, certain questions will need to be resolved.

3. State Liability in Cyberspace: Questions to Be Considered

This section provides an initial sketch of the main conceptual questions that will need to be addressed to develop a principle of state liability for transboundary harm in cyberspace: What constitutes “significant harm”? How should causation be evaluated? What duties does a state owe (or should owe) other states in cyberspace? What should standard of liability should be applied?\textsuperscript{169}

a. What Constitutes Significant Harm?

Because this definition relies on “significant harm” as a limiting factor, much will depend on what level and kinds of harm are “significant.”\textsuperscript{170}

\textsuperscript{168} Michael N. Schmitt, \textit{Grey Zones in the International Law of Cyberspace}, 42 \textit{Yale J. Int’l L. Online} 1, 2 n.11 (2017) (discussing legal restrictions on covert countermeasures, including the requirement of prior or subsequent notice); \textit{see also} \textit{Tallinn Manual} 2.0, \textit{supra} note 41, at 120 (Rule 21) cmt. 10; \textit{see also} id. at 120 (Rule 21) cmt. 12 (noting situations where notice may not be required).

\textsuperscript{169} More instrumental questions—such as how this legal regime should be developed and enforced—are considered in Part IV.

\textsuperscript{170} Many conventions refer to “significant,” “serious,” or “substantial” harm or damage
Thousands of damaging cyberoperations occur on a daily basis—what level of economic injury constitutes an international cybertort? A thousand dollars’ worth of damage? A million? Ten million? Is this an objective assessment, or will the understanding of what is significant depend on the wealth of the attacked entity? Should more abstract harms—such as interference in an election—be recognized and addressed? Domestic tort law certainly recognizes a wide variety of harms, including violations of property or constitutional rights as well as physical, emotional, and reputational injuries. There might be a similar panoply of more abstract harms at the international level. The U.S. Department of Defense’s Cyber Strategy, for example, showcases the potential breadth of the concept. It states that the DoD has an obligation to “defend the United States and its interests against cyberattacks of significant consequence,” regardless of whether they constitute cyberwarfare. While noting that cyberincidents will be “assessed on a case-by-case and fact-specific basis,” it declares that “significant consequences may include loss of life, significant damage to property, serious adverse U.S. foreign policy consequences, or serious economic impact on the United States.”

As with many other questions of evaluating and valuing tort violations, these are questions of fact that are best left to the “jury”—which, in the international legal order, is comprised of the community of states. States, like plaintiffs in domestic law, will determine what injuries they will absorb and which are worth challenging; other states’ responses to such accusations will be instrumental in developing norms about what constitutes significant harm. Indeed, as is often the case in international technological regulation,
the inherent ambiguity of “significant harm” is a strength: it is a tech-neutral standard that permits coherent but flexible legal development.177

b. What Duties Do States Owe Other States?

Much depends on the question of what duties a state owes (or should owe) other states. Any claim that a state is liable for transboundary harm associated with a cyberoperation must first demonstrate that states have (1) a duty not to cause transboundary harm in cyberspace; (2) a duty to prevent or mitigate the causation of transboundary harm in cyberspace, which might also be characterized as a duty of due diligence; and/or (3) a duty to compensate for transboundary harm caused by their cyberoperations or that occurs due to a lack of due diligence. These three conceptions of state duties might be understood as interrelated or completely independent.

A duty not to cause transboundary harm is fairly self-explanatory: states would have a duty not to engage in any activities that result in transboundary harm. A duty not to cause transboundary harm is distinct from a duty to prevent the causation of transboundary harm; the former would apply only to state and state-sponsored activities, while the latter is a more general charge to monitor and limit what other states and non-state actors do on its territory or within its jurisdiction or control. Under the former conception, North Korea would have had a duty not to engage in potentially harmful cyberoperations against the United States; under the latter, it would have had a duty to prevent other states or non-state actors from deploying anything from within its jurisdiction or control to engage in such activities.

The second option—a duty to prevent the causation of transboundary harm—might also be characterized as a duty of due diligence, which has some precedent as a freestanding duty in international law. In the Pulp Mills case, the ICJ found that all states have

[an obligation to act with due diligence in respect of all activities which take place under the jurisdiction and control of each party. It is an obligation which entails not only the adoption of appropriate rules and measures, but also a certain level of vigilance in their enforcement and the exercise of administrative control applicable to public and private operators . . . to safeguard the rights of the other party.]178

More and more, scholars are discussing due diligence as an independent


standard for evaluating appropriate state action in cyberspace. Schmitt argues that states should shoulder additional due diligence obligations in cyberspace, given that they have a “due diligence obligation with respect to both government and private cyber infrastructure on, and cyber activities emanating from, their territory.” The Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations has two rules regarding due diligence, recognizing it as a general principle and applying it in the cyber context.

Due diligence standard might span a spectrum of requirements. What might be characterized as a “strong” due diligence standard—a general duty to prevent others from causing transboundary harm—carries the risk of creating an incentive for states to exercise complete control over information technologies. This is exactly what the Shanghai Cooperation Organization states have been advocating, and exactly what Western states have been resisting in the interest of preserving freedom of expression and the free exchange of information on the internet. As Jack Goldsmith has noted, while questioning the likelihood of a cybersecurity treaty,

The United States might one day be willing to accept comprehensive U.S. government monitoring and 24/7 real-time police or military pursuit in the private network in exchange for a serious clamp-down on malicious activity from Russia and China. But the idea is unthinkable today. . . . What we need to do to protect ourselves in the cyber realm is in deep conflict with our commitments to limited government and private control of the communications infrastructure.


180 Rule 6 provides, “A State must exercise due diligence in not allowing its territory, or territory or cyber infrastructure under its governmental control, to be used for cyber operations that affect the rights of, and produce serious adverse consequences for, other States.” TALLINN MANUAL 2.0, supra note 41, at 30 (Rule 6). Rule 7 states, “The principle of due diligence requires a State to take all measures that are feasible in the circumstances to put an end to cyber operations that affect a right of, and produce serious adverse consequences for, other States.” Id. at 43 (Rule 7).

181 Cf. JACK GOLDSMITH, CYBERSECURITY TREATIES: A SKEPTICAL VIEW 9 (2011) (“The United States might one day be willing to accept comprehensive U.S. government monitoring and 24/7 real-time police or military pursuit in the private network in exchange for a serious clamp-down on malicious activity from Russia and China. But the idea is unthinkable today. . . . What we need to do to protect ourselves in the cyber realm is in deep conflict with our commitments to limited government and private control of the communications infrastructure.”).
In contrast, a “weak” due diligence duty would arise only with regard to known harms, states should be expected to take only reasonably feasible measures to minimize those harms, and there should be no duty to monitor networks and no duty of prevention.\(^\text{182}\)

For various political reasons, there is recent state practice that could be understood as supporting a duty of due diligence. In early 2017, a variant of the WannaCry ransomware “crippled 200,000 computers in more than 150 countries,”\(^\text{183}\) making it the largest ransomware assault to date.\(^\text{184}\) The malware employed a hacking tool called “Eternal Blue,” which was originally developed by the U.S. National Security Agency (NSA) and subsequently leaked by a hacking group called the “Shadow Brokers.”\(^\text{185}\) While there is growing evidence that the attacks are linked to North Korean hackers,\(^\text{186}\) China has blamed the United States for the attack, presumably on the rationale that the United States had allowed a dangerous tool to fall into the wrong hands or had not taken sufficient action to minimize the harm associated with its leaked tool.\(^\text{187}\) The NotPetya malware attacks also used Eternal Blue, resulting in similar calls for NSA to “help the rest of the world defend against the weapons it created.”\(^\text{188}\)

Finally, rather than there being a duty to not cause transboundary harm or a duty of due diligence, states may have a duty to compensate victims for any harms caused by their cyberoperations and/or caused by inadequate due diligence.\(^\text{189}\) This could be understood as a natural corollary of the other two potential duties, where the failure to fulfill the other duties triggers a duty to

\(^{182}\) Cf. TALLINN MANUAL 2.0, supra note 41, at 40-50 (Rules 6, 7); see also Scott J. Shackelford, Scott Russell & Andreas Kuehn, Unpacking the International Law on Cybersecurity Due Diligence: Lessons from the Public and Private Sectors, 17 CHI. J. INT’L L. 1 (2016).

\(^{183}\) Goldman, supra note 119.

\(^{184}\) Perlroth & Sanger, supra note 118.

\(^{185}\) Goldman, supra note 119.

\(^{186}\) Nicole Perlroth, More Evidence Points to North Korea in Ransomware Attack, N.Y. TIMES, May 22, 2017, at B4 (including the fact that the “WannaCry attacks used the same command-and-control server used in the North Korean hack of Sony Pictures Entertainment in 2014”).

\(^{187}\) Paul Mozur & Jane Perlez, China is Reluctant to Blame North Korea, Its Ally, for Cyberattack, N.Y. TIMES, May 17, 2017, at A8 (“Despite evidence suggests a North Korean role in the ransomware attack, the most common reaction among experts and on Chinese social media was to blame the United States.”).

\(^{188}\) Perlroth, Scott & Frenkel, supra note 114.

\(^{189}\) In other words, a duty not to cause transboundary harm might be understood as a liability rule. See Guido Calabresi & A. Douglas Melamed, Property Rules, Liability Rules, and Inalienability: One View of the Cathedral, 85 HARV. L. REV. 1089 (1972).
compensate. Alternatively, there may be a general presumption that states may engage in activities not prohibited by international law even if doing so causes transboundary harm, provided that the state subsequently compensates victims for any associated harms. The *Lotus* case—which held that state actions not expressly prohibited under international law are permitted—would support this more limited conception, and given state interest in preserving room to play, this is the most likely candidate for general adoption.

c. How Should Causation Be Evaluated?

A second limiting factor will be whether a given state can be determined to have caused an international cybertorts. Causation is hardly a simple legal concept, but it is at least a familiar one with well-traced complications. And, as in domestic tort law, establishing whether a given act caused a given harm will necessarily be a fact-specific analysis.

Causation of significant harm (the liability standard) is fundamentally different from attribution of an internationally wrongful act (the state responsibility standard): “[I]nternational liability of a State associated with its obligation not to cause harm to other States requires no attributability of the act to the State. . . . International liability arises out of injurious consequences which, according to the natural law of causation, must result from activities over which the State has or should have direct or indirect control or that lie within its jurisdiction.”

Given this, it seems likely that “causation” will be interpreted far more broadly than “attribution” with regard to the activities of non-state actors—especially if states are understood as having a due diligence obligation.

Eric Talbot Jenson and Sean Watts have suggested that a duty of due diligence may help mitigate the attribution problem. Imagine a scenario where State A is the victim of a malicious cyberoperation conducted by State B, which is routed through State C’s cyber infrastructure. While State A can determine the cyberoperation came from State C, it cannot reasonably identify the original perpetrator. Absent a due diligence duty, State A could

---

190 S.S. Lotus (Fr. v. Turk.), 1927 P.C.I.J. (ser. A) No. 10 (Sept. 7) (observing that, because of the consensual natural of the international legal order, “Restrictions upon the independence of States cannot be therefore be presumed”); *but see* Unilateral Declaration of Independence in Respect of Kosovo, Advisory Opinion, 2010 I.C.J. 141 ¶ 8 (July 22) (declaration of Judge Simma) (2010) (criticizing the famous *Lotus* dictum as an outdated, “excessively deferential” approach).


192 The state responsibility attribution standard for the activities of non-state actors is extremely circumscribed. *See infra* Part II.B.2.

193 Jenson & Watts, *supra* note 72.
not hold States B or C responsible nor engage in countermeasures against either; with a due diligence duty, State A could inform State C of the harm and, should State C fail to take reasonable actions to end it, State C would have committed an internationally wrongful act permitting State A to engage in countermeasures against State C.\footnote{194} However, as Jenson and Watts recognize, this potential benefit comes with attendant costs.\footnote{195} This conception of due diligence expands both when and against whom a victim state might use countermeasures, risking increased conflict escalation.\footnote{196}

But if due diligence is understood as expanding the number of states that could be held liable for compensation, as opposed to responsible for an internationally wrongful act justifying the use of countermeasures, some of the problems Jenson and Watts identify with expanding attribution would be minimized. Assigning liability, rather than responsibility, for due diligence violations might reduce the proliferation of self-help\footnote{197} and make multilateral solutions more attractive.\footnote{198}

d. What Standard of Liability Should Apply?

Finally, what standards of liability should be applied in evaluating state liability for their international cybertorts? As in domestic tort law, there are good reasons to employ different standards for different levels of intent, as there are fundamental differences between unforeseeable accidental damage, likely accidental damage, non-accidental damage, and intentional damage.\footnote{199}

Certainly, a state should be held liable for intended harms and from harms resulting from its ultrahazardous activities. Intentional torts and ultrahazardous activities—those that involve a risk of “significant

\footnote{194} Id. at 1567-68.
\footnote{195} Id. at 1568 (observing that a due diligence standard might contribute to the “erosion of State internalization of international law, proliferation of resorts to self-help, hindrance of multilateral and collective capacity, and faulty assignments of culpability”).
\footnote{196} Id. at 1577 (“In short, by presenting more opportunities for more States to allege more breaches of international law, due diligence potentially increases the frequency of States’ resort to countermeasures and their accompanying potentially destabilizing effects.”).
\footnote{197} Id. at 1573-74.
\footnote{198} Id. at 1574-75. Admittedly, however, this substitution of state liability for state responsibility does little to address that fact that this is “a proxy approach,” whereby the perpetrator can completely evade consequences, id. at 1575, and it may even exacerbate the problem of rule erosion, to the extent it encourages states to engage in “efficient breaches,” id. at 1568-73.
\footnote{199} Whether a state should be held vicariously liable for the actions of non-state actors or other states operating on its territory or employing devices within its jurisdiction or control will depend first on whether a duty to prevent the causation of transboundary harm is established.
transboundary harm that is either unforeseeable or, if foreseeable, unpreventable even if a state takes due care”—are almost always evaluated under a strict liability standard.200

States should also be held at least partially liable for unintended harms resulting from their not-unlawful activities, though the appropriate standard of liability is less obvious. Some suggest that “liability of a State may be said to be strict or almost absolute, regardless of fault, intention or negligence, for activities within its jurisdiction or on a sea-going vessel or spacecraft carrying its flag or registered in its territory.”201 A strict liability standard simplifies the analysis: if it can be determined that a state’s action or inaction caused transboundary harm, it will be liable for the costs associated with that harm. There are also arguments for a lesser standard of liability for harms resulting from negligence or from socially-useful activities. Walton, for example, argues that “due diligence” is best understood as a standard of liability, rather than as a freestanding independent duty,202 and that it provides the appropriate standard when evaluating unintentional harms associated with socially-useful activities.

Oren Gross has also argued that the state victim to a harmful cyberoperation should bear some liability for failures to take appropriate cybersecurity measures.203 A victim state’s particularly egregious cybersecurity practices might be treated as a kind of contributory or comparative negligence that mitigates another state’s liability for its international cybertorts.

* * * * *

A cyberoperation like the Sony hack does not fit squarely into the transnational cybercrime nor the cyberwarfare categories. Instead, it is conceptually and legally useful to identify the Sony hack as an international cybertort. Shifting to a tort-law framework invites a more thoughtful consideration of how the principle of state liability for transboundary harms could apply in cyberspace and its attendant benefits.

201 See XUE HANQIN, TRANSBOUNDARY DAMAGE IN INTERNATIONAL LAW 14 (2003); Sucharitkul, supra note 135, at 835.
202 Walton, supra note 17, at 1497 (“[I]f due diligence is the appropriate standard by which to judge state conduct at the level of low-intensity cyber attacks, then such an approach would have to recognize the underlying duty to prevent transboundary harm—given that this is the only primary duty that governs the low-intensity space.”).
203 Gross, supra note XX.
However, there is another aspect to cyberoperations like the DNC hack worth discussing: in addition to being cyberespionage that cost the DNC hefty sums, the action was likely also intended to sow confusion in a matter of internal importance and possibly even alter the outcome of a U.S. presidential election. While the DNC hack may not have been itself unlawful, imagine if Russian actors instead hacked voting machines and altered individual votes. If such an action caused significant harm, it might be an international cybertort. But it would also be something more—unlawful interference.\footnote{Egan, supra note 166 (“[A] cyber operation by a State that interferes with another country’s ability to hold an election or that manipulates another country’s election results would be a clear violation of the rule of non-intervention.”).}

III. STATE RESPONSIBILITY FOR INTERNATIONALLY WRONGFUL ACTS

In contrast to the principle of state liability for transboundary harm, the law of state responsibility aims to hold states accountable for their internationally wrongful acts. After a brief review of the law of state responsibility, this Part discusses two kinds of unlawful interference—violations of state sovereignty and intervention—that would ordinarily trigger the applicability of state responsibility and why cyberspace facilitates such activities. It concludes, somewhat counterintuitively, that instead of expanding existing definitions of internationally wrongful acts to cover these cyber-enabled interferences, states should use the possibility of state liability to deter such cyberoperations.

A. The Law of State Responsibility

The law of state responsibility is intended to create accountability mechanisms for states that engage in any “internationally wrongful act,” defined as “conduct consisting of an action or omission” that “constitutes a breach of an international obligation of the State” and “is attributable to the State under international law.”\footnote{Draft Articles, supra note 42, art. 2.} If a state is responsible for an internationally wrongful act, it is obligated to make full reparation.\footnote{Id. art. 31.}

1. Breach of an International Obligation

From its inception, the ILC’s focus in codifying the law of state responsibility was limited to the topic of wrongful acts. According to the
Draft Articles, “The essence of an internationally wrongful act lies in the non-conformity of the State’s actual conduct with the conduct it ought to have adopted in order to comply with a particular international obligation.”\textsuperscript{207} The conduct a state “ought to have adopted” might be found in customary international law, treaty law, or general principles of the international legal order.\textsuperscript{208} However, an act is not a breach of an international obligation “unless the State is bound by the obligation in question at the time the act occurs.”\textsuperscript{209} The number of potentially internationally wrongful acts a state can engage in is limited only by its international obligations.\textsuperscript{210}

It is worth noting that if the law evolves such that states are understood to have a duty to compensate those harmed by their international cybertorts, the failure to provide compensation might itself constitute an internationally wrongful act triggering the applicability of the law of state responsibility and its broader remedial measures.

2. Attribution

“Attribution,” in the state responsibility context, “denote[s] the operation of attaching a given action or omission to a State.”\textsuperscript{211} Certainly, the actions of state organs are attributable to a state.\textsuperscript{212} States may be held responsible both for the actions of those non-state actors that are \textit{de facto} state organs,\textsuperscript{213} as well as for the actions of non-state actors acting “on the instructions of, or under the direction or control of” a state in carrying out an operation.\textsuperscript{214}

The standard for determining when attribution for \textit{de facto} state organs is...
appropriate remains unresolved. The International Court of Justice (ICJ) has adopted a “strict control” test, while the International Criminal Tribunal for the Former Yugoslavia (ICTY) has employed a relatively relaxed “overall control” standard.\textsuperscript{215} If a non-state actor is a \textit{de jure} or \textit{de facto} state organ, the state will be responsible for all of its actions, regardless of whether they are \textit{ultra vires}.\textsuperscript{216}

Additionally, the acts of non-state actors may also be attributable to a state under Article 8 of the Draft Articles, which holds that “\textit{[t]he conduct of a person or group of persons shall be considered an act of a State under international law if the person or group of persons is in fact acting on the instructions of, or under the direction or control of, that State in carrying out the conduct.}\textsuperscript{217} In the Draft Articles, the ILC adopted the ICJ’s “effective control” test for Article 8 attribution.\textsuperscript{218} Under this standard, a state will only be held responsible for actions that occur in the context of an operation over which it exercises effective control,\textsuperscript{219} and it will only be responsible for a non-state actor’s \textit{ultra vires} actions that are “an integral part” of the operation.\textsuperscript{220}

The \textit{Tallinn Manual 2.0} ties state responsibility for a non-state actor’s cyberoperation to the ICJ’s “effective control” test under Article 8.\textsuperscript{221} As many scholars have argued, however, this standard (and the other tests for attribution) may be inappropriately high for determining state-sponsored


\textsuperscript{216} Draft Articles, \textit{supra} note 42, art. 7; \textit{id.} art. 7 cmts. 1-8 (describing supporting state practice and judicial decisions); see also Tadić, Case No. IT-94-1-A, ¶ 121.

\textsuperscript{217} Draft Articles, \textit{supra} note 42, art. 8. The ICJ has recognized this as reflecting customary international law. \textit{Bosnian Genocide}, 2007 I.C.J. ¶ 398.

\textsuperscript{218} Draft Articles, \textit{supra} note 42, art. 8, cmts 3-5.


\textsuperscript{220} Draft Articles, \textit{supra} note 42, art. 8 cmt. 3.

\textsuperscript{221} \textit{TALLINN MANUAL 2.0}, \textit{supra} note 41, at 96 (Rule 17) cmt. 6.
cyberoperations carried out by non-state actors.\textsuperscript{222} Attributing the acts of non-state actors to states is never an easy undertaking, and it is complicated in cyberspace by the opportunities for anonymity and misdirection.

Ultimately, absolute certainty regarding attribution is rarely possible; instead, a state seeking to hold another responsible for an internationally wrongful act are expected to independently judge a variety of facts to make a reasonable determination as to whether there is justification for attribution.

3. Reparations

Once an internationally wrongful act is attributed to a state, the state is then “under an obligation to make full reparation for the injury caused by the internationally wrongful act.”\textsuperscript{223} The concept of reparation under the law of state responsibility is far broader than the compensation suggested by the principle of state liability. Reparation might “take the form of restitution, compensation and satisfaction, either singly or in combination.”\textsuperscript{224} Restitution requires “re-establish[ing] the situation which existed before the wrongful act was committed.”\textsuperscript{225} Monetary compensation is required to the extent damage is not made good by restitution.\textsuperscript{226} Satisfaction—which may entail acknowledging the breach, expressing regret, or a formal apology—is

---

\textsuperscript{222} See, e.g., Michael Gervais, Cyber Attacks and the Laws of War, 30 BERKELEY J. INT’L L. 525, 549-50 (2012) (arguing that a victim state may use force against a state that refuses to prevent malicious cyberoperations emanating from its territory); Catherine Lotrionte, State Sovereignty and Self-Defense in Cyberspace: A Normative Framework for Balancing Legal Rights, 26 EMORY INT’L L. REV. 825, 890 (2012) (arguing that a victim state should be able to use force against states that are “directly or indirectly” involved in a non-state actor’s cyberoperations); Peter Margulies, Sovereignty and Cyber Attacks: Technology’s Challenge to the Law of State Responsibility, 14 MELBOURNE J. INT’L L. 496 (2013) (proposing a “virtual control” standard, which would “impos[e] responsibility on a state that has provided financial or other assistance to private groups” and shift the burden of proof to the accused state); Matthew J. Sklerov, Solving the Dilemma of State Responses to Cyberattacks: A Justification for the Use of Active Defenses against States Who Neglect Their Duty to Prevent, 201 MILITARY L. REV. 1 (2009) (arguing that states may use force against third-party states who do not take sufficient precautions against their servers being used for cyberoperations). But see Michael N. Schmitt & Liis Vihul, Proxy Wars in Cyber Space: The Evolving International Law of Attribution, 1 FLETCHER SECURITY REV., no. 2, at 65 (2014) (suggesting that the existing standards will remain high).

\textsuperscript{223} Draft Articles, supra note 42, art. 31. The 2005 Basic Principles expand this list to include rehabilitation and guarantees of nonrepetition. Basic Principles, supra note 82, ¶ 18; see also Draft Articles, supra note 42, art. 30(b) (imposing an obligation on a state responsible for an internationally wrongful act to offer guarantees of non-repetition).

\textsuperscript{224} Draft Articles, supra note 42, art. 34.

\textsuperscript{225} Id. art. 35.

\textsuperscript{226} Id. art. 36.
required to the extent the damage cannot be made good by restitution or compensation.\textsuperscript{227} The appropriate form of restitution will depend on the kind and scope of the harm, and of course, full reparation may not always be possible.

\textbf{B. Cyber-Facilitated Interference}

States regularly attempt to influence other states’ actions in myriad ways—through economic aid and sanctions, propaganda, political maneuvering, and shows of military force. International law permits many such influences, but recognizes two kinds of interference—violations of state sovereignty and intervention—as internationally wrongful acts.\textsuperscript{228} When attributable to a state, such unlawful interferences trigger the law of state responsibility.

While these concepts are well-established in principle, their scope is often unclear. Furthermore, given how cyberspace facilitates interference and that states are reluctant to term such activities internationally wrongful acts, the line between lawful and unlawful interference is becoming further blurred.

1. Unlawful Interference: Violations of State Sovereignty and Interventions

State sovereignty is one of the foundational concepts of the international legal order. As articulated by Max Huber in the 1928 Island of Palmas arbitral award, “Sovereignty in the relations between States signifies independence. Independence in regard to a portion of the globe is the right to exercise therein, to the exclusion of any other State, the functions of a State.”\textsuperscript{229} The U.N. Charter consecrates the concept, grounding its legitimacy “on the principle of the sovereign equality of all its Members.”\textsuperscript{230}

The customary prohibition on intervention forbids “all States or groups

\textsuperscript{227} Id. art. 37. Interest may be necessary to ensure full reparation. Id. art. 38.

\textsuperscript{228} While the terms “interference” and “intervention” are sometimes used interchangeably, it is useful to distinguish between them. Interference encompasses both lawful and unlawful meddling; intervention is coercive and therefore prohibited. \textit{See} 1 \textsc{Oppenheim`S International Law} 432 (Sir Robert Jennings & Sir Arthur Watts eds., 9th ed., 2008) [hereinafter \textsc{Oppenheim}].

\textsuperscript{229} Island of Palmas (Neth. v. U.S.) 2 R.I.A.A. 829, 838 (Perm. Ct. Arb. 1928). The \textit{Tallinn Manual} 2.0 conceives of state sovereignty as having an internal and external component and, by extension, divides potential violations of state sovereignty into two categories: “(1) the degree of infringement upon the target State’s territorial integrity; and (2) whether there has been an interference with or usurpation of inherently governmental functions.” \textit{Tallinn Manual} 2.0, \textit{supra} note 41, at 13 (Rule 2); \textit{id.} at 16 (Rule 3); \textit{id.} at 20 (Rule 4) cmt. 10.

\textsuperscript{230} U.N. Charter art. 2, para. 1.
of States to intervene directly or indirectly in the internal or external affairs of other States.”\textsuperscript{231} This prohibition is well-established in international law, with some even going so far as to consider it \textit{jus cogens}.\textsuperscript{232} It was first codified in a multilateral treaty in the 1933 Montevideo Convention: “No state has the right to intervene in the internal or external affairs of another.”\textsuperscript{233} Although the prohibition on intervention is not specifically mentioned in the U.N. Charter,\textsuperscript{234} post-Charter institutions have regularly reaffirmed it\textsuperscript{235}; the ILC noted it in its draft articles on the rights and duties of states\textsuperscript{236}; the U.N. General Assembly has issued a number of resolutions reiterating it\textsuperscript{237}; and the International Court of Justice (ICJ) regularly acknowledges it.\textsuperscript{238}

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{231} Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, ¶ 205 (June 27); see also id. ¶ 209 (holding that, where interference takes the form of a use or threat of force, Article 2(4) and the customary norm of non-intervention are coterminous).
\item \textsuperscript{232} Watts, supra note 82, at n.27-28 (citing sources).
\item \textsuperscript{233} Montevideo Convention on the Rights and Duties of States art. 8, Dec. 26, 1933, 49 Stat. 3097, T.S. 881.
\item \textsuperscript{234} While the principle of state sovereignty can be read to include the principle of non-intervention, as a formal matter the U.N. Charter only explicitly prohibits intervention by itself or other U.N. bodies. U.N. Charter art. 2, para. 7 (“Nothing contained in the present Charter shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state or shall require the Members to submit such matters to settlement under the present Charter; but this principle shall not prejudice the application of enforcement measures under Chapter VII.”).
\item \textsuperscript{235} See, \textit{e.g.}, Charter of the Organization of American States art. 19, Feb. 27, 1967, 33 I.L.M. 1004 (“No State or group of States has the right to intervene, directly or indirectly, for any reason whatever, in the internal or external affairs of any other State. The foregoing principle prohibits not only armed force but also any other form of interference or attempted threat against the personality of the State or against its political, economic, and cultural elements.”).
\item \textsuperscript{236} International Law Commission, Draft Declaration on Rights and Duties of States with Commentaries art. 3 (1949) (“Every State has the duty to refrain from intervention in the internal or external affairs of any other State.”).
\end{enumerate}
\end{footnotesize}
The prohibition on intervention can be understood as flowing directly from the concept of state sovereignty: if states have a right to the independent, exclusive exercise of state functions, other states are necessarily prohibited from taking coercive actions that would impair that right.\textsuperscript{239} According to one reading, the prohibition on intervention can be understood as protecting the non-territorial, “metaphysical aspect of sovereign integrity rather than its physical dimension.”\textsuperscript{240}

Alternatively, violations of sovereignty and intervention can be understood as separate categories of internationally wrongful acts that, while clearly related, do not completely overlap.\textsuperscript{241} Certainly, there are acts that could be considered sovereignty violations that do not seem to meet the standard for interventions: there is a strong argument that the Sony hack constituted a violation of U.S. sovereignty, although it was not sufficient coercive regarding a state’s affairs to qualify as an intervention.\textsuperscript{242} It is harder, however, to conceive of an intervention that would not also violate a state’s sovereignty, given that coercive interference in a state’s affairs would necessarily constitute a significant “interference with or usurpation of inherently governmental functions.”\textsuperscript{243}

2. An Elusive Line Between Lawful and Unlawful Interference

Both state sovereignty and the prohibition on intervention are well-established in principle, but the scope of their application resists clear codification—in part because adjudications of these issues tend to be fact-specific and resist generalizations, in part because experts are divided on whether it is appropriate to apply older concepts to new kinds of technologically-facilitated interference.

a. State Sovereignty

The difficulty in defining the scope of what constitutes a violation of state sovereignty is illustrated by the Tallinn Manual 2.0 experts’ inability to agree on its borders. Most of the experts agreed that “cyber operations constitute a

\textsuperscript{239} Cf. Nicaragua, 1986 I.C.J. ¶ 205 (stating that the prohibition on intervention forbids states from meddling in “matters in which each State is permitted, by the principles of State sovereignty, to decide freely”).

\textsuperscript{240} See, e.g., Russell Buchanan, The International Legal Regulation of State-Sponsored Cyber Espionage, in INTERNATIONAL CYBER NORMS: LEGAL, POLICY & INDUSTRY PERSPECTIVES 65, 78 (2016).

\textsuperscript{241} See, e.g., Schmitt & Vihul, supra note 248.

\textsuperscript{242} Schmitt, supra note 8.

\textsuperscript{243} TALLINN MANUAL 2.0, supra note 41, at20 (Rule 4) cmt. 10.
violation of sovereignty in the event they result in physical damage or injury, as in the case of malware that causes the malfunctioning of the cooling elements of equipment, thereby leading to overheating that results in components melting down” and that “the causation of physical consequences by remote means on that territory likewise constitutes a violation of sovereignty.” 244 However, the experts could not reach consensus on the question of whether “a cyber operation that results in neither physical damage nor the loss of functionality amounts to a violation of sovereignty.” 245

Similarly, there are competing arguments regarding whether the DNC hack would constitute a violation of U.S. sovereignty. Assuming that it can be attributed to Russia, some would characterize it as a sovereignty violation because it involved nonconsensual intrusion into U.S. cyberinfrastructure. 246 However, there is a minority but growing viewpoint that “mere compromises or thefts of data are not violations of sovereignty, but rather routine facets of espionage and competition among States.” 247

Further complicating matters, senior U.S. officials have recently argued there is no rule against violations of sovereignty in international law—rather, there is a principle that state sovereignty is to be respected that takes different forms in different forums, and that the rules for cyber are still in flux. 248

b. Intervention

The prohibition on intervention has also resisted clear delineation. 249 First, it is not obvious what state activities are shielded. The ICJ has stated

244 Id. at 20 (Rule 4) cmt. 11.
245 Id. at 21 (Rule 4) cmt. 14.
246 Watts, supra note 10.
247 Id. Regarding this, Ryan Goodman has made an interesting and somewhat counterintuitive point on this: if the misappropriation and distribution of information associated with the DNC hack is not a violation of international law—if it is not a violation of state sovereignty or intervention—the practice could be employed unilaterally as a punitive retorsion. Ryan Goodman, International Law and the US Response to Russian Election Interference, JUST SECURITY (Jan. 5, 2017, 8:01 AM), https://www.justsecurity.org/35999/international-law-response-russian-election-interference/.
248 For an argument in favor of this understanding, see Gary P. Corn & Robert Taylor, Sovereignty in the Age of Cyber, AM. J. INT’L. L. UBOUND (forthcoming 2017). Gary Corn is the Staff Judge Advocate of the U.S. Cyber Command, and Robert Taylor is the former Principal Deputy General Counsel of the U.S. Department of Defense. For a responsive critique, see Michael N. Schmitt & Liis Vihul, Respect for Sovereignty in Cyberspace, 95 TEX. L. REV. 1639 (2017).
249 CHATHAM HOUSE, THE PRINCIPLE OF NON-INTERVENTION IN CONTEMPORARY INTERNATIONAL LAW 3, 6 (2007).
that protected state affairs include “choice of a political, economic, social, and cultural system, and the formulation of foreign policy”\textsuperscript{250}; according to a more recent Chatham House report, prohibited activities might also include, depending on the circumstances, “[i]nterference in political activities,” “[s]upport for secession,” and “[s]eeking to overthrow the government—so-called ‘regime change.’”\textsuperscript{251} The concept of the \textit{domaine réservé} helps describe what state activities are protected from intervention, but the “displacement of a matter or issue from the \textit{domaine réservé} constitutes neither an overall eradication or waiver of the principle of non-interference, nor an open season on influencing conditions in another State’s territory.”\textsuperscript{252} Ultimately, there is no definitive list of what state affairs are and are not protected.\textsuperscript{253}

Second, lawful interference is often distinguished from prohibited intervention based on the degree of coercion exercised.\textsuperscript{254} But what is coercion? Oppenheim defines unlawful intervention as that which is “forcible or dictatorial, or otherwise coercive, in effect depriving the state intervened against of control over the matter in question.”\textsuperscript{255} Dispatching armed forces to another states will certainly constitute prohibited intervention; it is less clear is whether and when economic, political, and psychological pressures are sufficiently coercive to satisfy the legal requirement for prohibited intervention.\textsuperscript{256}

Confusion in the doctrine is leading some states to exploit the grey areas, and others to err on the side of repression. Consider Ecuador’s reaction to the DNC hack. WikiLeaks founder Julian Assange has been living in exile in Ecuador’s London embassy since 2012, avoiding a Swedish rape investigation which he believes to be the cover story for an American extradition. With the expressed intent of hobbling WikiLeaks’ interference in the 2016 U.S. general election and to evade any hint of responsibility for

\footnotesize{\textsuperscript{250} Nicaragua, 1986 I.C.J. ¶ 205.}
\footnotesize{\textsuperscript{251} CHATHAM HOUSE REPORT, supra note 249, at 7.}
\footnotesize{\textsuperscript{252} Watts, supra note 82.}
\footnotesize{\textsuperscript{253} Id.}
\footnotesize{\textsuperscript{254} For example, the ICJ identified “coercion” as the element “which defines, and indeed forms the very essence of, prohibited intervention.” Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 I.C.J. 14, ¶ 205 (June 27).}
\footnotesize{\textsuperscript{255} OPPENHEIM, supra note 228, at 432; see also TALLINN MANUAL 2.0, supra note 41, at 319 (Rule 66), cmt. 21 (“The key is that the coercive act must have the potential for compelling the target State to engage in an action that it would otherwise not take (or refrain from taking an action it would otherwise take).”).}
\footnotesize{\textsuperscript{256} There is a minority view that Article 2(4) prohibits political and economic coercion; the majority understanding is that it prohibits only threats or uses of force. Hathaway et al., supra note 67, at 842.}
facilitating unlawful interventions, Ecuador cut Assange’s internet access.\textsuperscript{257} However, the controversial leaked emails were originally obtained by Russian hackers, not Assange or Ecuador\textsuperscript{258}—and while their disclosure likely constituted disfavored interference, it probably was not sufficiently coercive to meet the definition for unlawful intervention.

Given these ambiguities, a determination of whether a given act constitutes a prohibited violation of state sovereignty or unlawful intervention necessitates a fact-specific inquiry, focused on the existence, validity, and scope of state consent; the degree and kind of coercive activity; and whether any coercive acts by a non-state actor can be attributed to another state.\textsuperscript{259} Unfortunately, cyberspace facilitates interference while simultaneously confusing the facts relevant to categorization.

3. Increased Likelihood of Interference

The U.S. attribution of the DNC hack to Russia, reignited an ongoing conversation regarding whether and how foreign actors may use new technologies in an attempt to influence elections.\textsuperscript{260} Such cyber-enabled acts might range from publicizing hacked private communications to disseminating misinformation to exploiting voting machine vulnerabilities to manipulating social media.\textsuperscript{261} Nor are these pure hypotheticals: states have long used cyberoperations to influence elections,\textsuperscript{262} and U.S. security experts...
and non-governmental organizations have recently identified areas of the electoral infrastructure that are particularly vulnerable to tampering.\(^{263}\)

Election manipulation is just one example of how cyberspace permits an entirely new level of non-physical but nonetheless pervasive interference. Meanwhile, not only are some of the traditional obstacles to interference inapplicable in cyberspace, many of the existing deterrents are less effective.

First, the shift from physical space to the cyber realm enables states to engage in entirely new levels of invasive but non-violent interference. States can reach into the very heart of another state’s operations and steal, manipulate, or delete critical information, allowing them to obtain or compromise information on a scale previously unimaginable.

Meanwhile, many of the traditional practical obstacles to different kinds of interference are simply not applicable. Historically, influence operations might require extensive intelligence, personnel, or military resources, costs that limited which states could intervene and how often they were willing to do so. That is no longer the case. States can now engage in all kinds of invasive operations without any individual ever crossing a border, and at dramatically lower price points. In 2014, for example, industry experts estimated that it would cost roughly about $10,000 for a state to develop Stuxnet-like malware.\(^{264}\) More recently, an expert calculated the cost of developing and operating a new Advanced Persistent Threat—malware that can “break into any [specific] target, exfiltrate data, analyse it, and produce intelligence product”—for one year to be a mere $2 million.\(^{265}\)

By lowering their cost, cyber lowers the barrier to entry, increasing the number of states able to engage in such interference. Certainly, electoral interference is nothing new. Dov Levin estimates that, from 1946 to 2000, either the United States or the U.S.S.R./Russia interfered in another country’s elections 117 times, which roughly translates to an intervention in “one out of every nine competitive national level executive elections” during this fifty-

---


year period. Nor is the fact that powerful states are interfering in each other’s elections particularly new. What is new is that the United States and Russia are no longer the only states capable of such interference. Small states and even non-state actors may now have the resources and capacity to interfere in the affairs of powerful states.

Not only is it easier for more states to engage in more invasive cyberoperations, but traditional legal deterrents are less effective. Most physical violations or sovereignty and interventions are public. The victim state can respond immediately, and outside states, international organizations, non-governmental organizations, and non-state actors can observe the original action, the victim state’s response, and react accordingly. Invasive cyberoperations, in contrast, can be simultaneously pervasive, destructive, and entirely secret. How is a victim state supposed to react when it does not know the perpetrator, the meaning of the act—or even that the act occurred? Finally, even if the victim state identifies the act and can reasonably attribute it to another state, it has few lawful responsive options, resulting in the state paralysis noted above.

In short, the inherent characteristics of cyberspace undermine many of the practical and legal deterrents to interference while simultaneously promising greater payoffs. The clear implication is that cyber-enabled interference is likely to skyrocket—raising the question of how victim states should respond.

C. How State Liability Might Minimize Resort to Countermeasures

Certainly, some cyber-enabled interferences will easily meet the traditional definitions for violations of state sovereignty or intervention and can be addressed under the existing law of state responsibility. But states are grappling with how to respond to hacks, info dumps, and other new forms of interference occurring on an unprecedented, cyber-enabled scale. For example, in the wake of an increase in “aggressive cyberespionage” targeting German politicians and a November 2016 cyberincident that caused 900,000 Germans to lose internet access, the head of Germany’s foreign intelligence

---

266 Dov Levin, When the Great Power Gets A Vote: The Effects of Great Power Electoral Interventions on Election Results, 60 INT’L STUD. Q. 189 (2016); see also William J. Daugherty, Executive Secrets: Covert Action and the Presidency (2006) (discussing specific examples of U.S. covert actions to influence another state’s elections).


268 See supra Part I(B).
service warned that Russia might be interfering in Germany’s elections. Former CIA Acting Director Michael Moore made similar statements regarding Russian interference in the 2016 U.S. general election: “It is an attack on our very democracy. It’s an attack on who we are as a people. A foreign government messing around in our elections is . . . . the political equivalent of 9/11.”

Recent cyberoperations have raised one of the perennial questions associated with new technology: is it enough to apply the existing rules, or is there something unique about the traits or effects of the new technology that requires new law? When considering how best to conceive of sovereignty and intervention in the cyber context, it quickly becomes clear that the existing prohibitions, standing alone, do not sufficiently address problems that arise in the cyber domain. For example, the DNC hack alone might not be considered a violation of U.S. sovereignty because nothing was damaged, nor a prohibited intervention because the dissemination of private hacked emails information was not sufficiently coercive.

And yet, there is an understandable desire to stretch existing terms regarding internationally wrongful acts to these new activities, both to clarify their wrongfulness and justify the use of countermeasures. Some are arguing for expanding the definition of coercion to encompass activities like the DNC hack, either by altering or completely doing away with the coercion requirement. Duncan Hollis suggests that the intentional nature of the leak and the timing “clearly sought to maximize attention (and corresponding impacts) on the U.S. domestic political campaign process,” and Schmitt

270 Id.
271 Morell & Kelly, supra note 12.
argues that the “sounder” understanding of the DNC hack is that it was coercive because “the cyber operations manipulated the process of elections and therefore caused them to unfold in a way that they otherwise would not have.”

However, lowering or eliminating the coercion requirement would transmute many of today’s routine and minor interferences into prohibited interventions—with undesirable side effects. Not only might a lowered threshold deter some entities from engaging in humanitarian activities, actions that states are currently expected to let go unpunished in the interest of preserving international peace would become internationally wrongful acts, justifying state resort to escalatory unilateral countermeasures.

Similar problems would attend a more expansive definition of state sovereignty. As Walton has noted, “[A] definition of sovereignty that is too broad might inadvertently cover a whole host of cross-border intrusions accepted in an interconnected world, such as the extraterritorial effects of a state’s telecommunications, industrial, monetary, and environmental activities.”

Nor would expanding the universe of what constitutes unlawful interference necessarily create more options for a victim state. As noted above, practical and legal limitations on the use of countermeasures in response to cyberoperations strictly curtail their utility.

Instead of expanding the already-ambiguous scope of unlawful interferences to cover new kinds of invasive cyberoperations, victim states might instead take the intermediate— but possibly more effective—option of bringing a state liability claim for associated injuries. Although such a claim may not be able to address all of the harms associated with a politically-motivated interference, it will allow the victim state to name and shame the perpetrator and possibly recover compensation without the risk of creating problematic precedent, encouraging conflict escalation, or becoming itself responsible for an internationally wrongful act.

275 Schmitt, supra note 168, at 8.
276 Hathaway, supra note 45, at 49.
277 Walton, supra note 17, at 1477.
278 See supra Part I(B).
279 Should international law evolve to encompass a more liberal understanding of the coercion element for cyber-enabled interventions, there is good reason to limit any such development to cyberspace. It is worth reiterating that the scope of the prohibition on intervention in the physical world is still unclear and evolving; mixing in new practices developed in cyberspace risks further muddying the waters. The fact that rules developed in the physical space do not apply well in cyberspace suggests that the reverse might be true. To avoid creating inappropriate precedent, cyber-enabled unlawful interferences with no physical effects should be distinguished from physical interferences. Doing so allows for the
The prior two Parts have drawn on fundamental principles from tort law and international law to construct more a comprehensive regime of state accountability for different kinds of harmful actions in cyberspace, grounded in both state liability for acts with injurious consequences and state responsibility for internationally wrongful acts. The next Part considers how best to develop these accountability regimes.

IV. A COMPREHENSIVE SYSTEM OF STATE ACCOUNTABILITY IN CYBERSPACE

It is worth noting at the outset that this Article’s proposed categories and associated accountability regimes could be immediately incorporated within the existing international enforcement mechanisms. States could label harmful cyberoperations international cybertorts and demand compensation through formal or informal channels.

That being acknowledged, it would be far preferable if states create an independent institution with the expertise and investigative resources to impartially assess state accountability in cyberspace, the flexibility to adapt to changing technologies, and the enforcement authority to deter victim states from engaging in inappropriate and potentially escalatory self-help. An institution would also contribute to the considered and comprehensive development of the international law of cyberspace, as its determinations would bridge the gap between positivist treaty law and the unhurried development of a customary international law of cyberspace.

A. State Interest in Developing the Law

Notwithstanding differing opinions on how best to do so, most agree that states must play a central role in developing the law of cyberspace. To this end, states have produced and published domestic cyber policies, engaged in confidence building measures, signed bilateral and multilateral non-development of a cyber-specific countermeasures regime under the law of state responsibility, without impacting the equilibrium struck by the U.N. Charter and existing law of countermeasures.


binding political agreements regarding state behavior in cyberspace, and negotiated and ratified the Convention on Cybercrime. But while states have an interest in clarifying the rules of the information superhighway, they do not want well-enforced speed limits. The trick will be finding a middle ground that maintains some leeway to speed while avoiding widespread crashes and pileups.

States have myriad reasons to be interested in participating in the development of the law of cyberspace. In 2015, cyber threats were identified as the international community’s top security threat, and President Obama declared the threat of cyber warfare a national emergency. Aside from the obvious national security implications, developing the law of cyberspace is vital to growing the global digital economy. In listing problems associated with the lack of shared peacetime norms of state behavior in cyberspace, U.S. cybersecurity experts on the President’s Commission on Enhancing National Cybersecurity noted that “the international digital economy lacks the coherent systems necessary to effectively address cross-border malicious cyber activity. . . . The void in technical, policy, and legal conventions hampers information sharing and interoperability, and creates an opening for criminals to launch attacks and conduct other malicious cyber activity.” Given this, “[c]oordinated and effective international harmonization and cooperation are needed in order to realize the full economic promise of the

282 Convention on Cybercrime, supra note 9.
283 The 2015 U.S. Law of War Manual, for example, has been critiqued for doing little to expand upon the public record of the U.S. understanding of the law of cyberspace, despite professing an interest in elucidation. Sean Watts, Cyber Law Development and the United States Law of War Manual, in INTERNATIONAL CYBER NORMS: LEGAL, POLICY & INDUSTRY PERSPECTIVES 49, 63 (Anna-Maria Osula & Henry Roigas eds., 2016) (“More than simply confirmation of persistent ambiguities in the operation of the law of war in cyberspace, the ambiguities the Manual leaves unresolved are strong evidence of the US’ comfort with these uncertainties and legal voids. . . . [T]he Manual indicates significant state reticence toward and even a present inclination against definitive clarity and precision in this challenging domain of state competition.”).
284 In the context of international cybertorts, much of this leeway will be preserved in the term “significant harm,” see supra Part II.B, and by limiting state duties in cyberspace to compensation for the injuries caused by their cyberoperations, see supra Part II.C.2.
287 COMMISSION ON ENHANCING NATIONAL CYBERSECURITY, REPORT ON SECURING AND GROWING THE DIGITAL ECONOMY 47 (Dec. 2016).
288 Id.
nation and the world, and to allow for the efficient flow of information and ideas.”\footnote{Id. at 47. In the interest of ensuring “an open, fair, competitive, and secure global digital economy,” these experts recommend that the United States “encourage and actively coordinate with the international community in creating and harmonizing cybersecurity policies and practice and common international agreements on cybersecurity law and global norms of behavior.” Id.}

As Kristen Eichensehr has observed, “Even for those who may be skeptical of international engagement and international law or norms in general, the Commission’s [economic-based] perception that international coordination is crucial should be persuasive.”\footnote{Kristen Eichensehr, \textit{The Economic Incentives for International Cybersecurity Coordination}, \textit{JUST SECURITY} (Dec. 6, 2016; 12:09 PM), https://www.justsecurity.org/35310/economic-incentives-international-cybersecurity-coordination/.
}

There have been some initial steps towards the development of international cyberspace peacetime norms.\footnote{The United States has been an active participant in this process. It originally worked to establish what is now the “global affirmation of the applicability of international law to state behavior in cyberspace,” and it is currently attempting to foster “international consensus on additional norms and principles of responsible state behavior in cyberspace that apply during peacetime.” U.S. DEP’T STATE POLICY, \textit{supra} note 281, at 12-13. These include four priority norms: (1) “[A] State should not conduct or knowingly support cyber-enabled theft of intellectual property, trade secrets, or other confidential business information with the intent of providing competitive advantages to its companies or commercial sectors”; (2) “[A] State should not conduct or knowingly support online activity that intentionally damages critical infrastructure or otherwise impairs the use of critical infrastructure to provide service to the public”; (3) “[A] State should not conduct or knowingly support activity intended to prevent national computer security incident response teams (CSIRTs) from responding to cyber incidents” and “should not use CSIRTs to enable online activity that is intended to do harm”; and (4) “[A] State should cooperate, in a manner consistent with its domestic and international obligations, with requests for assistance from other States in investigating cyber crimes, collecting electronic evidence, and mitigating malicious cyber activity emanating from its territory.” Egan, \textit{supra} note 166.

The U.S. approach to developing international cybersecurity norms has been a study in what former U.S. Legal Advisor Harold Koh has called “Twenty-First-Century International Lawmaking”—namely, a combination of “nonlegal understandings,” “layered cooperation,” and “diplomatic law talk.” Harold Hongju Koh, \textit{Address: Twenty-First-Century International Lawmaking}, 101 GEO. LJ. ONLINE 1, 13-16 (2012).

to self-defense in cyberspace in a consensus report, as well as the applicability of the law of armed conflict’s principles of humanity, necessity, proportionality, and distinction.\textsuperscript{293} And while the 2017 GGE could not agree on a final draft report—due largely to disagreement about one draft paragraph—they made other important progress towards developing the cyberspace norms and principles.\textsuperscript{294}

States are also articulating norms in the process of exploring the utility of cybersecurity confidence-building measures (CBMs) and through unilateral pronouncements.\textsuperscript{295} CBMs help minimize arms races and conflict escalation by reducing uncertainty about other states’ capabilities. Proposed cyber CBMs tend to focus on information sharing; facilitating communication among stakeholders; and potential future international and domestic actions, such as commitments “to refrain from a certain activity of concern.”\textsuperscript{296}

\begin{flushleft}


\textsuperscript{295} In December 2013, for example, the Permanent Council of the Organization for Security and Cooperation in Europe established eleven CBMs in cyberspace; in March 2016, the Permanent Council expanded this list to sixteen. Organization for Security and Cooperation in Europe, Permanent Council Decision No. 1202, OSCE Confidence-Building Measures to Reduce the Risks of Conflict Stemming from the Use of Information and Communication Technologies (Mar. 10, 2016). The United States has also endorsed the use of CBMs. U.S. DEP’T STATE POLICY, supra note 281, at 15 (stating that “cyber CBMs have the potential to contribute substantially to international cyber stability” and proposing cyber-appropriate CBMs).

\textsuperscript{296} U.S. DEP’T STATE POLICY, supra note 281, at 15; see also Jack Goldsmith, Contrarian Thoughts on Russia and the Presidential Election, LAWFARE (Jan. 10, 2017), https://www.lawfareblog.com/contrarian-thoughts-russia-and-presidential-election (arguing that technologically advanced states should be more open to “cutting a deal”: joining agreements of mutual restraint, where they pledge to forego engaging in certain actions or activities in cyberspace).

There are significant obstacles, however, to creating effective confidence-building measures in cyberspace. First, there is the usual obstacle to CBMs: States are disinclined to share information about their capabilities. The United Kingdom and United States, for example, have been reluctant to share information about their offensive cyber capabilities with their NATO allies. David E. Sanger, As Russian Hackers Probe, NATO Has No Clear Cyberwar Strategy, N.Y. TIMES, Jun. 16, 2016, at A13. Second, this is another situation where practices developed in the physical world don’t translate well to the cyber realm. CBMs, which originated in the disarmament context, are usually state-based and depend on monitoring and verification mechanisms. However, cyberspace is dominated by non-state actors, and the “[a]nonymity, complexity, the intangible nature of digital systems, and the lack of knowledge about the intended use of hardware and software make any verification
Additionally, states are building consensus around norms by publicizing their domestic policies and individual legal assessments of high-profile cyber incidents.\textsuperscript{297} Brian Egan, former U.S. Legal Advisor to the Department of State, is one of many who has called on states to “publicly state their views on how existing international law applies to State conduct in cyberspace to the greatest extent possible in international and domestic forums,” which “will give rise to more settled expectations of State behavior and thereby contribute to greater predictability and stability in cyberspace.”\textsuperscript{298}

Non-state entities are playing a pivotal role in spurring an international conversation on these issues. One particularly influential project is the original Tallinn Manual on the International Law Applicable to Cyber Warfare and Tallinn Manual 2.0. Although they were the product of an initiative of the NATO Cooperative Cyber Defence Centre of Excellence, the \textit{Manuals} are not an official NATO project nor were they intended to reflect states’ views. Rather, they are a published collection of international law experts’ joint reasoning and determinations regarding permissible state action in cyberspace, and while they are formally nonbinding, many of their conclusions have been widely accepted as authoritative.\textsuperscript{299} Where states, civil society, or scholars have disagreed with particular conclusions, the \textit{Manuals} have sparked broader and more informed discussions.

\textbf{B. Existing Implementation Mechanisms}

States could promote this Article’s proposals unilaterally, by explicitly grounding retorsions or countermeasures on the fact that another state’s action constitutes an international cybertort or unlawful interference. However, not only are self-help regimes inherently unstable, they are simultaneously less effective and more escalatory in cyberspace.\textsuperscript{300}

\begin{footnotesize}

\textsuperscript{298} Egan, \textit{supra} note 166; see also Ben Buchanan & Michael Sulmeyer, Hacking Chads: The Motivations, Threats, and Effects of Electoral Insecurity 18 (Oct. 2016) (“[T]he United States should put forth a declaratory policy on the vital importance of elections, vowing to impose costs on any state that interferes with the integrity of the process.”).

\textsuperscript{299} As a result, the Tallinn Manual is often presented in non-cyber contexts as an example of how non-state entities can have particularly influential impact on the development of international law. See, e.g., Kenneth Anderson, Daniel Reisner & Matthew Waxman, \textit{Adapting the Law of Armed Conflict to Autonomous Weapon Systems}, 90 Int’l L. Stud. 386, 407-08 (2014); Croofof, \textit{The Killer Robots Are Here}, \textit{supra} note 52, at 1902.

\textsuperscript{300} See \textit{supra} Part I.B.
\end{footnotesize}
It would be far preferable for states to respond to harmful or intrusive cyberoperations—and thereby develop the relevant law—through institutional action. Compared with self-help measures, institutional responses are less escalatory and more legitimate.

Unfortunately, existing institutional responses are more difficult to orchestrate and less likely to achieve a desired result. A state victim to cyber-enabled interventions could petition the United Nations for collective sanctions or the Security Council for an authorization for a limited use of force against the perpetrator, but such petitions are unlikely to garner much support, in no small part because the main perpetrators of harmful cyberoperations are permanent members of the Security Council. Alternatively, a state victim to a harmful cyberoperation might file an ICJ suit alleging transboundary harm, violation of sovereignty, or intervention. But while the ICJ is well versed in international law, it lacks technical expertise in evaluating cyberoperations. It also has significant jurisdictional issues. The ICJ only has jurisdiction in contentious cases on the basis of state consent: states may agree to bring a specific issue before the Court by submitting a compromis, or states may accept the Court’s jurisdiction as generally compulsory. But many powerful states have refused to accept or have withdrawn from the ICJ’s compulsory jurisdiction, and alleged perpetrators are unlikely to agree to submit a compromis.

Given these political and practical limitations, states often resort to self-help measures—or do nothing. But what if there was a more appropriate institutional option?

C. A New Institution

Ideally, states would create a new, independent institution with the expertise and investigative resources to impartially assess state accountability in cyberspace. This entity could be charged with fact-finding; alternatively or additionally, it might be tasked with determining state liability or responsibility for cyberoperations and granted the authority to recommend appropriate reparations to deter victim states from engaging in inappropriate and escalatory self-help.

302 Id. art. 36(2).
Determining the author of a cyberoperation is technically difficult and requires skilled forensic analysis. When compared with the alternatives—individual states, state coalitions, and the ICJ—an independent institution will be better able to recruit and retain individuals with the necessary expertise, to conduct an unbiased investigation, and to make broadly credible findings.\textsuperscript{304} Delegating forensic tasks to an independent institution might also reduce disparities between states with different levels of domestic technological capabilities.

As there will rarely be direct evidence linking an entity to a cyberoperation or linking non-state actors to states, there will be vexing evidentiary issues to address in any case alleging state involvement in harmful cyberoperations. Both the victim state and the accused state will likely be unwilling to provide the access and information needed for an outside institution to fully perform the fact-finding function. In many situations, the accused state will have exclusive access to critical evidence proving or disproving its connection to the cyberoperation, but regardless of whether it sponsored the act, it will likely be reluctant to produce such evidence for national security reasons. However, this is not an entirely new problem for international tribunals: the ICJ has developed a process for dealing with such situations that a new institution could adapt as needed.\textsuperscript{305}

\textsuperscript{304} See also CLARKE & KNake, supra note 93, at 251 (proposing an “International Cyber Forensics and Compliance Staff,” an international organization with inspection teams to determine the origins of attacks, with the ability to place traffic monitoring equipment inside domestic networks).

\textsuperscript{305} In international tribunals generally, the party alleging a fact has the burden of proving it. Pulp Mills on the River Uruguay (Argentina v. Uruguay), Judgment, 2010 I.C.J. Rep. 14, ¶ 162 (Apr. 20). Circumstantial evidence is generally permissible (although it is often critically examined). See Michael P. Scharf & Margaux Day, The International Court of Justice’s Treatment of Circumstantial Evidence and Adverse Inferences, 13 Chi. J. Int’l L. 123, 147 (2012) (analyzing jurisprudence from the ICJ, the Permanent Court of Arbitration, the Eritrea-Ethiopia Claims Commission, and the NAFTA Claims Tribunal). In situations where a claim depends on evidence in the sole possession of the accused state, the ICJ has sometimes held that the burden of proof shifts to that state. Ahmadou Sadio (Guinea v. Democratic Republic of the Congo), Judgment, Merits 2012 I.C.J. Rep. 324, ¶ 55 (June 19); Gangaram Panday v. Suriname, Inter-American Court of Human Rights, Judgment, Merits, Series C No. 16, ¶ 49 (Jan. 21, 1994). More often, rather than shifting the burden of proof or making formal adverse findings of fact, see Eritrea-Ethiopia Claims Commission, Partial Award, Central Front, Eritrea’s Claims 2, 4, 6, 7, 8 & 22, (Apr 28, 2004) (reading negative inferences of fact against a state for failing to produce evidence), the ICJ has instead used nonproduction of evidence “as a license to resort liberally to circumstantial evidence where direct evidence would otherwise be preferred,” Scharf & Day, supra at 128. In its 1949 Corfu Channel decision, the ICJ determined that, in cases where key evidence was in the possession of the accused state, the accusing state would enjoy “a more liberal recourse to inferences of fact and circumstantial evidence,” provided there was no room for reasonable doubt.
And the state bringing a claim and alleging improper conduct might be more willing to provide supporting evidence to gain the legitimacy that would attend an independent assessment of that information.

An independent institution might also be granted the power to issue binding decisions regarding appropriate reparations for significantly harmful or intrusive cyberoperations. It might even assign punitive damages, both to sanction past violations and serve as a non-escalatory deterrent, or proscribe permissible individual or collective countermeasures. This would fit well with the existing regime of state responsibility, which charges states to explore resolving disputes in tribunals where possible and to refrain from engaging in countermeasures while a dispute is pending. Also, it might avoid the jurisdictional problems of the ICJ: states invested in developing the law of cyberspace but concerned about waiving their sovereign immunity more broadly might be more willing to waive their immunity to suit and accept the limited jurisdiction of a specialized tribunal.

The creation of an independent international institution with specialized investigative, adjudicative, and norm-building capabilities is hardly a revolutionary suggestion. The International Atomic Energy Agency is a well-respected organization that investigates and verifies state usage of lawful nuclear technologies. The American-Mexican Claims Commission, the U.N. Compensation Commission, the Iran-United States Claims Tribunal, and even the World Trade Organization might all be considered precedents. These and similar entities are often created to deter states from engaging in self-help, minimize the coordination issues of collective action, solve the jurisdictional problems of other existing institutional options, and proscribe appropriate sanctions.

Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. 43 (Feb. 26). In 2007, the Court revisited this evidentiary problem in the Bosnian Genocide case. It relied on circumstantial evidence to reach a legal conclusion regarding Serbia’s failure to prevent atrocities, but disregarded such evidence with regard to the claim that Serbia intended to commit genocide. This may be because, with regard to the latter claim, Serbia submitted direct evidence in support of its defense that it did not meet the intent requirement for the crime of genocide. Scharf & Day, supra at 143. As a general rule, the ICJ “will permit liberal reliance on circumstantial evidence so long as two conditions are met: (1) the direct evidence is under the exclusive control of the opposing party; and (2) the circumstantial evidence does not contradict any available direct evidence or accepted facts.” Id. at 131.

306 Relatedly, international investment tribunals are increasingly contributing to the development of relevant customary international law, with the full support of litigating states. See W. Michael Reisman, Canute Confronts the Tide: State Versus Tribunals and the Evolution of the Minimum Standard in Customary International Law, 30 ICSID REV. 616 (2015).
D. A Preferable Means of Law Creation

In the course of evaluating claims, a new institution would necessarily have to address novel questions of law, such as what duties states owe other states in cyberspace. This approach to creating a law of cyberspace is far preferable to awaiting the development of a law of cyberspace via the more traditional sources of international law, namely a treaty on state accountability in cyberspace or customary cyber international law. Both treaty law and customary international law are ill-suited to developing state accountability for cyberoperations, underscoring the utility of creating an independent institution.

1. The Unlikelihood of a Comprehensive Cybersecurity Treaty

Many consider treaties—written agreements between two or more states—to be the gold standard of international law. In contrast to other sources of international law, treaties are written documents describing legal rights and obligations of state parties to which states explicitly consent to be bound. Given this backdrop presumption, many hope that building international consensus around norms of state behavior in cyberspace may eventually lead to the codification of these norms in a broad, multilateral cybersecurity treaty.

For a variety of reasons, however, a constitutive cybersecurity treaty may not be possible. At the most basic level, there are few cyber-related subjects that permit mutually beneficial deals for states with differing technological capabilities, differing vulnerabilities, and differing beliefs about the appropriate amount of governmental control over the internet or the dangers posed by free speech.

Indeed, not only do states desire to regulate different activities in cyberspace, many states see other’s proposed norms as being antithetical to their own concerns. Goldsmith has observed, “[T]here are deep and fundamental clashes not only over what practices should be outlawed but also and more broadly over what the problem is.” As an example, he discusses

---

307 See supra Part II.C.2.
312 GOLDSMITH, supra note 181, at 4.
how the United States is promoting a norm against attacking civilian targets in part because it would disproportionately serve U.S. interests, given U.S. dependence on civilian networks, poor cybersecurity practices, and the fact that it already rarely attack other states’ civilian networks. Meanwhile, not only are Chinese civilian networks more secure than those in the United States, the Chinese military is not nearly as dependent on them. Why would China give up this potential military advantage in support of a norm against targeting civilians, Goldsmith questions, without gaining anything in return? He concludes that “[t]he distributional consequences of any such agreement may be such that some nations will be willing to risk the threats to infrastructures from non-cooperation because threats fall asymmetrically on their adversaries.”

Meanwhile, China has repeatedly attempted and failed to garner widespread support for its proposals recognizing states’ “cyber sovereignty”—the concept that state sovereignty justifies multilateral internet governance—because this is commonly perceived to be at odds with Western visions of internet freedom and U.S. interests in preserving multistakeholder internet governance.

There are also significant enforcement issues: Even assuming that states manage to negotiate a broad cybersecurity treaty with relatively narrowly-tailored terms that limit opportunities for creative interpretations, how will state compliance with those terms be verified? In short, a constitutive cybersecurity treaty is unlikely to be negotiated—and if one is, it is not likely to be effective.

Not only might it be impossible to negotiate or monitor an effective constitutive cybersecurity treaty, it is probably not an ideal means of developing the international law of cyberspace (contrary to what I and my co-authors have argued in the past). One of the primary strengths of a constitutive multilateral treaty is its stability, which justifies state investment during the negotiation and drafting process. However, all treaties risk becoming outdated as times and norms change—and law regulating new

313 Id. at 6.
314 GOLDSMITH, supra note 181, at 6-7 (“[T]he cybersecurity context is and will remain bedeviled by two types of definitional difficulty. The first arises from the nature of the activity itself, which makes precise definitions of weapons, effects, and targets difficult. . . . [Second, w]hen nations disagree sharply over the matter to be regulated, they tend to agree (if at all) in vague generalities that are not terribly useful for fostering true cooperation.”); see also Croofof, Killer Robots Are Here, supra note 52, at 1888-89 (discussing the importance of clear and narrowly tailored prohibitions in the effectiveness of a regulatory treaty).
315 Limited or bilateral cybersecurity treaties, however, may well be useful in some contexts, such as in the development of confidence-building measures.
316 Hathaway et al., supra note 67, at 880-84.
technology is particularly susceptible to early obsolescence. Instead of defaulting to the presumption that treaty law is superior, it is worth considering the relative benefits of other sources of international law.

2. The Difficulty with Developing Customary International Cyber Law

A rule of customary international law is recognized as existing when states generally engage in specific actions (the “state practice” element) and accept that those actions are obligatory or permitted (the “opinio juris sive necessitatis” element). In short, “a rule of customary international law is authoritative because states generally abide by it in the belief that it is law.”

While customary international law has been critiqued for lacking the clarity of written law, that indefiniteness is a strength—it is flexible and responsive to change, especially technological change. Accordingly, it may be preferable that the international law of cyberspace be grounded in customary international law rather than a constitutive treaty.

However, there is one significant drawback to awaiting the development of customary international cyber law: namely, cyber-specific customary international law is unlikely to develop organically in the near future. Evidence of state practice is a fundamental requirement to the formation of customary international law. Although cyber-based technologies foster the speedy development of customary international law generally by facilitating the dissemination of information and increasing the number of state

---

317 Crootof, Change Without Consent, supra note 109, at 242. While scholars, practitioners, and judges tend to favor the lex scripta of treaty law over customary international law for various functional reasons, as a matter of doctrine the two sources of international law are co-equal. Id. at 285 n.274 (citing sources).

318 It is important to distinguish between simply applying existing customary international law to cyberspace and identifying the development of cyber-specific customary international law. The former considers norms developed in the physical world and attempts to determine how they operate in cyberspace; the latter would examine norms that develop based on state practice in cyberspace.

functions, state practice in cyberspace is largely hidden. There simply aren’t enough examples to establish that states reliably act in a certain way in the belief that those actions are permitted or required by law. The few examples of state practice that come to light are the one-offs and the mistakes, and it would be foolish to ground a governance regime on sporadic, exceptional actions.

3. The Benefits of Institutional Legal Development

An institution charged with developing the law of cyberspace bridges the gap between difficult-to-obtain positivist treaty law and the unhurried development of a customary international cyber law. It is a Goldilocks solution: institutional decisions and reports will have the authority and clarity of written law while maintaining flexibility and responsiveness to changing technological capabilities.

An independent institution could also contribute to the proactive development of a customary international cyber law. The institutional process will force states to articulate their understandings of relevant legal obligations, which in turn will foster scholarly and practitioner debates. Additionally, by promulgating codes of conduct or best practices, an institution could increase the likelihood that the law of cyberspace develops in a cohesive, flexible manner.

CONCLUSION

New technology often exposes gaps and vagueness in existing law and undermines foundational assumptions and justifications of legal regimes—and cyberspace is no exception. Quite the contrary. Cyberspace is a particularly bewildering arena: its infrastructure is shared by civilians and militaries, governments and businesses; cyberoperations occur and must be rebuffed at super-human speeds; non-state actors can be equally—if not more—powerful than some states; and it can be difficult to identify transgressors, both because the source of cyberoperations can be masked and because states often operate through non-state actors. As a result, there is substantial normative confusion, as legal rules made for the physical world map do not always map well onto the cyber domain. Given this confusion,

---

320 Croots, Change Without Consent, supra note 109, at 245-47.
322 Cf. Wolff Heintschel von Heinegg, Territorial Sovereignty and Neutrality in Cyberspace, 89 INT’L L. STUD. 123 (2013) (noting that states have responded to the legal
states have a vested interest in clarifying the international laws of cyberspace, both to know what actions they may lawfully take and how they may lawfully respond to other states’ actions.

This Article draws on tort law and international law principles to construct a comprehensive system of state accountability in cyberspace, where states are both liable for their lawful but harmful acts and responsible for their wrongful ones. Not only does recognizing international cybertorts and its attendant regime of state liability permit new means of managing the harms associated with data destruction, ransomware, and cyberespionage, it preserves a bounded grey zone for state experimentation in cyberspace, minimizes the likelihood that victim states will resort to escalatory responses, and increases the chance that those harmed by cyberoperations will be compensated.

confusion of cyberspace by attempting to apply laws developed for the physical world, with mixed success); Duncan B. Hollis, Re-Thinking the Boundaries of Law in Cyberspace: A Duty to Hack?, in CYBERWAR: LAW & ETHICS FOR VIRTUAL CONFLICTS, 1-4, 14-30 (J. Ohlin et al. eds., 2014) (discussing the benefits and problems with reasoning by analogy in cyberspace).
SOVEREIGNTY IN THE AGE OF CYBER

Gary P. Corn* and Robert Taylor†

Introduction

International law is a foundational pillar of the modern international order, and its applicability to both state and nonstate cyber activities is, by now, beyond question. However, owing to the unique and rapidly evolving nature of cyberspace, its ubiquitous interconnectivity, its lack of segregation between the private and public sectors, and its incompatibility with traditional concepts of geography, there are difficult and unresolved questions about exactly how international law applies to this domain. Chief among these is the question of the exact role that the principle of sovereignty plays in regulating states’ cyber activities.

The technological structure and global interconnectedness of cyberspace offers both state and nonstate actors a medium through which to operate against a broad array of targets, free from the physical constraints of geography and territorial boundaries. At an increasing rate, states are not just utilizing, but also aggressively exploiting, cyberspace as a novel means for engaging in traditional statecraft, including activities that advance national security such as espionage and low-cost, asymmetric offensive operations. Likewise, nonstate actors routinely use cyberspace to conduct harmful activities that threaten individuals and business interests, as well as nations. For example, ISIS uses the internet to command and control its operations, spread its toxic propaganda, recruit new members, and incite violence globally. And while ISIS may not yet possess the cyber sophistication or capabilities of certain nation-states, it has demonstrated the ability and willingness to engage in offensive cyber operations. The vulnerabilities inherent in the cyber domain threaten states’ basic governmental functions, territorial security, political independence, and economic well-being, as well as the individual security and well-being of their citizens.

Through both custom and treaty, international law establishes clear proscriptions against unlawful uses of force and prohibits certain interventions among states. While questions remain as to the specific scope and scale of cyber-generated effects that would violate these binding norms, the rules provide a reasonably clear framework for assessing the legality of state activities in cyberspace above these thresholds, including available response options

* Staff Judge Advocate, United States Cyber Command.
† Visiting Scholar, Harvard Law School Spring Term 2017, Former Principal Deputy General Counsel, Department of Defense. The views expressed are those of the authors and do not necessarily reflect the views of the United States Cyber Command, the Department of Defense, or the U.S. Government. Authors are listed in alphabetical order.

for states. Below these thresholds, there is insufficient evidence of either state practice or *opinio juris* to support assertions that the principle of sovereignty operates as an independent rule of customary international law that regulates states’ actions in cyberspace. The process of adapting the existing legal framework to cyberspace must reflect the unique aspects of the domain in order to ensure the law continues to serve the goal of maintaining peace and stability while enabling states to thwart existing threats and prevent new threats from emerging.

**Jus Ad Bellum and Nonintervention**

International law provides a framework for cooperation that is foundational to the successful preservation of international peace and security, and includes variations across different domains that account for the particularities of each domain. Both the *jus ad bellum*, as reflected in Article 2(4) of the UN Charter and customary international law (CIL), as well as the CIL rule of nonintervention, are well-recognized binding norms applicable to interstate relations. There is general consensus that the *jus ad bellum* applies fully to cyber activities that rise to the level of a use of force. There has been a great deal of discussion as to what cyber actions would actually amount to a use of force, and many commentators have speculated on that issue. Despite the lack of complete clarity, it is generally accepted that at a minimum, cyber activities that proximately result in death, injury, or significant destruction, or that represent an imminent threat thereof, constitute a use of force.

There is also general consensus that cyber actions that amount to a *prohibited intervention* also violate international law. This category would include actions that are coercive or dictatorial and bear on matters of sovereignty such as the choice of a political, economic, social, and cultural system, as well as the formulation of foreign policy. The quintessential example of a violation of the principle of nonintervention is one state coercively interfering in the internal political process of another state, such as by altering the votes recorded and thereby affecting the results of an election. Although the nonintervention rule is firmly enshrined in CIL, there is a dearth of state practice, let alone *opinio juris*, regarding the contours of its applicability to cyber activities. Further development of how the nonintervention rule applies to cyber activities is critical to informing policymakers on available response options such as the correlative doctrine of countermeasures.

**Activities Below the Level of Prohibited Intervention**

The *jus ad bellum* and the principle of nonintervention provide limited guidance in the realm of cyber because the vast majority of cyber operations are something less than a use of force, and do not fit squarely within the traditionally recognized elements of the nonintervention rule. Whether extant international law regulates this less-intrusive class of cyber activities is therefore a critical question. Some argue that limitations imposed by the concept of sovereignty fill this normative space—that sovereignty is itself a binding rule of international law that precludes virtually any action by one state in the territory of another that violates the domestic law of that other state, absent consent. However, law and state practice instead indicate that sovereignty serves as a principle of international law that guides state interactions, but is not itself a binding rule that dictates results under international law. While this principle of sovereignty, including territorial sovereignty, should factor into the conduct of every cyber operation, it does not establish an absolute bar against individual or collective state cyber operations that affect

---


cyberinfrastructure within another state, provided that the effects do not rise to the level of an unlawful use of force or an unlawful intervention.

Sovereignty is a fundamental principle of international law and is considered a “basic constitutional doctrine of the law of nations.” Since at least the Treaty of Westphalia and the creation of the modern nation-state, sovereignty has been understood to encompass two distinct but related aspects—internal and external sovereignty. As a general concept, it refers to “the collection of rights held by a state, first in its capacity as the entity entitled to exercise control over its territory and second in its capacity to act on the international plane, representing that territory and its people.” It is understood as encompassing “the whole body of rights and attributes which a state possesses in its territory, to the exclusion of all other states, and also in its relation with other states.” With respect to internal sovereignty, it signifies the independent right, in regard to a portion of the globe, “to exercise therein, to the exclusion of any other state, the functions of a state.” This well recognized, near-exclusive right over the domaine réservé of the state—those matters of governance and jurisdiction committed to the sole responsibility of the state and its official actors—is at the heart of, and specifically protected by, the principle of nonintervention. In addition, internal sovereignty allows a state to prohibit acts within or affecting its territory as an exercise of governmental authority. However, international law does not obligate other states to refrain from all activities that might infringe upon or operate to the prejudice of the territorial state’s internal sovereignty.

A perhaps more foundational aspect of sovereignty is the equality of states in the international order. Sometimes referred to as external sovereignty, this corollary principle to the exclusive authority of the state to exercise jurisdiction and governance within its territory refers to the recognition in the international order of the absolute equality and independence of all states. External sovereignty forms the unifying principle of international law—that only states, qua states, have the legal personality necessary to create and be bound by international law. The principle of sovereign equality underlies the well-recognized premise in international law that “[r]estrictions on the independence of states cannot … be presumed.” This premise, known as the Lotus rule, has long been understood to stand for the proposition that states are free to act on the international plane except to the extent that their actions are proscribed by treaty or customary international law.

Since the rise of the modern nation-state, countries have applied the doctrine of sovereignty in different ways, at times developing specific international law regimes tailored to the particular circumstances. For example, it is widely recognized that states have unquestioned authority to prohibit espionage within their territory under their domestic laws, but it is also widely recognized that international law does not prohibit espionage. States have long engaged in espionage operations that involve undisclosed entry and activities within the territory of other states, subject only to the risk of diplomatic consequences or the exercise of domestic jurisdiction over intelligence operatives if discovered and caught. Within this framework, it is understood that espionage may violate international law only when the modalities employed otherwise constitute a violation of a specific provision of

5 Id. at 448.
7 Island of Palmas (Neth. v. U.S.), 2 R.I.A.A 829, 838 (Perm. Ct. Arb. 1928). See also Samantha Besson, Sovereignty, in MAX PLANCK ENCYCLOPEDIA OF PUBLIC INTERNATIONAL LAW para. 119 (2011) (noting that sovereignty is generally characterized as the “powers and privileges resting on customary law which are independent of the particular consent of another state”).
8 CHATHAM HOUSE, supra note 3.
9 UN Charter art. 2(1) (“The Organization is based on the principle of the sovereign equality of all its Members.”).
10 Under limited circumstances, international organizations may create international law, but only to the extent states have conferred on them the authority to do so.
11 The S.S. Lotus (Fr. v. Turk.), 1927 P.C.I.J. (ser. A) No. 10, at 18 (Sept. 7).
international law, such as an unlawful intervention or a prohibited use of force. Thus states conduct intelligence activities in and through cyberspace, and generally, “to the extent that cyber operations resemble traditional intelligence and counter-intelligence activities ... such cyber operations would likely be treated similarly under international law.”12 This framework applies equally to cyber operations directed at terrorist cyber infrastructure located within the territory of another state.

Further, the differences in how sovereignty is reflected in international law with respect to the domains of space, air, and the seas further support the view that sovereignty is a principle, subject to adjustment depending on the domain and the practical imperatives of states rather than a hard and fast rule. For instance, in the case of the space domain, objects in orbit are beyond the territorial claims of any nation, and outer space—including outer space above another state’s territory—is available for exploitation by all. In the case of the air domain, the regime is highly restrictive, such that any unconsented entry into the airspace of another state is regarded as a serious violation of international law subject to such exceptions as self-defense, Security Council authorization, or force majeure. In the case of the seas, many entries into and transits through the territorial waters of another state are permissible without the consent of that state, but there are conditions under which such entry would be a violation of international law—it depends on the particular facts and circumstances. The fact that states have developed vastly different regimes to govern the air, space, and maritime domains underscores the fallacy of a universal international law.13

A contrary view asserts that sovereignty is more than a foundational principle of international law, upon which rules such as the prohibitions against the use of force and unlawful intervention are based. In this view, sovereignty is itself a rule of international law that can be violated.13 This view is frequently stated in terms of violations of or interferences with a state’s territorial sovereignty.14 However, this confuses the concept of territorial sovereignty, inherent in the notion of internal sovereignty discussed above, with the more precise concepts of territorial integrity and the inviolability of borders protected through Article 2(4), the Charter more broadly, and CIL. The proscriptions against violating territorial integrity or borders involve a threshold of harm much higher than the mere conduct of cyber operations limited to affecting, for example, cyber facilities owned or operated by terrorists, criminals, or by third states, located inside another state’s borders.15 Ultimately, whether and precisely when nonconsensual cyber operations below the threshold of a prohibited intervention violate international law is a question.

14 Some point to cases such as Costa Rica v. Nicaragua (Certain Activities Carried Out by Nicaragua in the Border Area (Costa Rica v. Nicar.) and Construction of a Road in Costa Rica along the San Juan River (Costa Rica v. Nicar.) (Dec. 16, 2015)), Corfu Channel (Corfu Channel, supra note 6), and Armed Activities on the Territory of the Congo (Armed Activities on the Territory of the Congo (New Application 2002) (Dem. Rep. Congo v. Uganda), 2006 ICJ REP. 6 (Feb. 3)) as support for this position. While it is true that the ICJ has referred to violations in those cases, in each instance the facts ruled on involved substantial military presence, de facto control of territory, and in some instances, violent operations, all of which implicate higher thresholds than the sovereignty-as-a-rule proponents assert. For example, in Armed Activities on the Territory of the Congo, the ICJ found that Uganda’s “unlawful military intervention” inside the territory of the Democratic Republic of the Congo (DRC) constituted a use of force in violation of Article 2(4) and a prohibited intervention, as well as military occupation of some DRC territory, and as such, constituted a violation of the DRC’s sovereignty and territorial integrity in the broader sense.
that must be resolved through the practice and opinio juris of states, developed over time and in response to the need of states effectively to defend themselves and provide security for their citizens.\(^\text{16}\)

The ramifications of the growing debate as to whether and how international law regulates state actions in cyberspace below the nonintervention threshold are easily observed in the case of the use of cyber capabilities in the fight against widely recognized terrorist organizations such as ISIS. In these cases, owing to the distributed nature of cyberspace, operations may involve cyber effects directed against terrorist-controlled infrastructure on networks or systems located in states outside areas of hostilities, and not subject to any preexisting authority to use force. While these “subintervention” cyber activities should consider the sovereignty of the states in whose territory these terrorist infrastructures reside, this does not answer the key question of whether or how sovereignty proscribes such cyber activities. While the principle of sovereignty should factor into the conduct of any cyber operation, it does not itself establish a bar against individual or collective state cyber operations against all cyber infrastructure within another state, particularly those controlled by widely-recognized terrorists and used for terrorist activity. In short, sovereignty is a principle, not a rule, and its legal consequences are not fully formed in this area.

**Applying this Approach to Terrorist Operations**

ISIS, like other transnational organizations, has a vast social media presence and uses the internet to communicate with its members and supporters, to recruit individuals to its cause, and to promote its views and activities. Its media operations are decentralized and originate from servers in numerous states throughout the world, encouraging crowdsourcing and volunteerism by individuals not directly associated with the terrorist organization. Additionally, ISIS adherents have formed the Cyber Caliphate and have taken active malicious cyber actions over the internet.\(^\text{17}\) This means that ISIS followers and adherents both inside and outside ISIS-controlled territory operate on servers and infrastructure scattered across the globe, taking advantage of the transparency and permeability of borders that characterize the internet. These states may have limited or no knowledge that ISIS is utilizing servers or cyber infrastructure under their sovereign authority. Further, these states may lack the capability to effectively counter or even discover ISIS’s cyber threat.

If the view were adopted that sovereignty is a rule violated by any action illegal under the domestic law of a state, states seeking to disrupt distributed terrorist cyber infrastructure would be under an obligation to either seek Security Council authorization or the consent of the state in whose territory the infrastructure resides. The nature of cyber operations and capabilities often require high degrees of operational security and the flexibility to act with speed and agility. Operating through a consent model could in important cases surrender operational initiative to the terrorist adversary or render response options unworkable. Further, these actions could involve cyber effects in, yet invisible to, the territorial state, but that only manifest operationally in the area of hostilities.

Because the doctrine of sovereignty does not prevent all actions by one state that affect another state or even “encroachment on other sovereign jurisdictions,”\(^\text{18}\) a state involved in operations against ISIS, such as the United States, is not precluded from taking action against ISIS’s cyber facilities in other states, even without the consent of the host state, unless doing so constitutes a prohibited intervention or use of force. Where the proposed cyber action is focused solely against the individual accounts or facilities of terrorists or terrorist organizations widely recognized as such, and when the cyber actions will generate only de minimis effects on nonterrorist infrastructure

---

\(^{16}\) Brian Egan (Legal Adviser, U.S. Dep’t of State), *International Law and Stability in Cyberspace* (November 10, 2016).


\(^{18}\) Egan, _supra_ note 16.
within the host state, international law does not preclude those cyber actions. States may choose, as a matter of policy, to seek the host state’s consent before taking such actions, but there is no customary law requirement to do so. The host state, exercising its sovereign domestic authorities, could proscribe such activities just as states proscribe espionage, but such proscription would be a matter of domestic law of the host state, rather than a matter of international law.

ISIS and other terrorist or criminal organizations, and indeed some states, are using and will continue to use the internet and cyber facilities to advance their malevolent ends. Effectively countering these activities is vital to national and international security. States must ensure that any response complies with international law and takes into account each state’s sovereignty. Properly understood and applied, international law and the doctrine of sovereignty do not preclude per se individual and collective state cyber actions directed against terrorist-owned or -operated cyber facilities, or other malevolent infrastructure, located within other states.
CONCLUDING OBSERVATIONS ON SOVEREIGNTY IN CYBERSPACE

Gary P. Corn* and Robert Taylor†

In *Sovereignty in Cyberspace: Lex Lata Vel Non?*, Michael Schmitt and Liis Vihul argue that territorial sovereignty is a primary rule of international law that limits cyber activities.¹ They recognize, however, that not all cyber effects constitute violations of territorial sovereignty, and like Rule 4 in the *Tallinn Manual 2.0* and its commentary, they acknowledge a distinct lack of consensus among the Tallinn participants on the critical question of applicable thresholds. Problematically, they do not identify the necessary state practice and *opinio juris* that would be required to establish either the primary rule that they proffer or the existence and contours of the exception they would recognize.

Schmitt and Vihul, as well as Phil Spector,² look to sources dealing with very different domains and very different kinds of activities, and attempt to divine a rule where we see an absence of binding law. In this regard, it is telling that at no point has the UN Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (UNGGE), the only international body to date charged with the task of examining how international law applies to cyber operations by states, identified sovereignty as a primary rule of international law that would, absent a justification, bar some or any nonconsensual cyber operations below the threshold of a prohibited intervention within the territory or on the infrastructure of another state. On the contrary, before backtracking and failing to reach a consensus report in 2017 on the applicability of international law to cyber operations vel non, the 2015 UNGGE adopted “general and declaratory” language that “[s]tate sovereignty and international norms and principles that flow from sovereignty apply to the conduct by States of [information and communications technologies]-related activities,” a position fully consonant with the preponderance of sources cited in Phil Spector’s essay.³ The 2015 UNGGE then went on to adopt a number of nonbinding, voluntary peacetime norms, many if not all of which would be superfluous if sovereignty were the primary rule that Schmitt, Vihul, and Spector assert it is.

Outside of the clearly established primary rules against the use of armed force and against unlawful intervention, it remains for states to consider the demarcation between what is lawful and what is not. How traditional principles and extant rules of international law apply to the emerging cyberspace domain are critical questions requiring the work and attention of the international community. That work must take into account the unique nature of cyberspace, as well as the foundational principle of respect for a nation’s sovereign authority, without sliding into the overly simplistic position that any prejudicial action in another state’s territory constitutes a breach of international law.

* Staff Judge Advocate, United States Cyber Command.
† Visiting Scholar, Harvard Law School Spring Term 2017. Former Principal Deputy General Counsel, Department of Defense. The views expressed are those of the authors and do not necessarily reflect the views of the United States Cyber Command, the Department of Defense, or the U.S. Government. Authors are listed in alphabetical order.

law whenever those actions might violate that state’s domestic law. It also must take into account the need of states to effectively defend against actors—state, nonstate, and hybrid actors—intent on causing harm from outside the borders of the defending state.
PANEL V:

LEGAL ETHICS AND LEAKING BY GOVERNMENT ATTORNEYS

DISCUSSANTS:
F. SCOTT KIEFF
SUSAN GIBSON
ABA MODEL RULES OF PROFESSIONAL CONDUCT

Client-Lawyer Relationship

RULE 1.6: CONFIDENTIALITY OF INFORMATION

(a) A lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent, the disclosure is impliedly authorized in order to carry out the representation or the disclosure is permitted by paragraph (b).

(b) A lawyer may reveal information relating to the representation of a client to the extent the lawyer reasonably believes necessary:

(1) to prevent reasonably certain death or substantial bodily harm;

(2) to prevent the client from committing a crime or fraud that is reasonably certain to result in substantial injury to the financial interests or property of another and in furtherance of which the client has used or is using the lawyer's services;

(3) to prevent, mitigate or rectify substantial injury to the financial interests or property of another that is reasonably certain to result or has resulted from the client's commission of a crime or fraud in furtherance of which the client has used the lawyer's services;

(4) to secure legal advice about the lawyer's compliance with these Rules;

(5) to establish a claim or defense on behalf of the lawyer in a controversy between the lawyer and the client, to establish a defense to a criminal charge or civil claim against the lawyer based upon conduct in which the client was involved, or to respond to allegations in any proceeding concerning the lawyer's representation of the client;

(6) to comply with other law or a court order; or

(7) to detect and resolve conflicts of interest arising from the lawyer’s change of employment or from changes in the composition or ownership of a firm, but only if the revealed information would not compromise the attorney-client privilege or otherwise prejudice the client.

(c) A lawyer shall make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client.

Comment

[1] This Rule governs the disclosure by a lawyer of information relating to the representation of a client during the lawyer's representation of the client. See Rule 1.18 for the lawyer's duties with respect to information provided to the lawyer by a prospective client, Rule 1.9(c)(2) for the lawyer's duty not to reveal information relating to the lawyer's prior
representation of a former client and Rules 1.8(b) and 1.9(c)(1) for the lawyer's duties with respect to the use of such information to the disadvantage of clients and former clients.

[2] A fundamental principle in the client-lawyer relationship is that, in the absence of the client's informed consent, the lawyer must not reveal information relating to the representation. See Rule 1.0(e) for the definition of informed consent. This contributes to the trust that is the hallmark of the client-lawyer relationship. The client is thereby encouraged to seek legal assistance and to communicate fully and frankly with the lawyer even as to embarrassing or legally damaging subject matter. The lawyer needs this information to represent the client effectively and, if necessary, to advise the client to refrain from wrongful conduct. Almost without exception, clients come to lawyers in order to determine their rights and what is, in the complex of laws and regulations, deemed to be legal and correct. Based upon experience, lawyers know that almost all clients follow the advice given, and the law is upheld.

[3] The principle of client-lawyer confidentiality is given effect by related bodies of law: the attorney-client privilege, the work product doctrine and the rule of confidentiality established in professional ethics. The attorney-client privilege and work product doctrine apply in judicial and other proceedings in which a lawyer may be called as a witness or otherwise required to produce evidence concerning a client. The rule of client-lawyer confidentiality applies in situations other than those where evidence is sought from the lawyer through compulsion of law. The confidentiality rule, for example, applies not only to matters communicated in confidence by the client but also to all information relating to the representation, whatever its source. A lawyer may not disclose such information except as authorized or required by the Rules of Professional Conduct or other law. See also Scope.

[4] Paragraph (a) prohibits a lawyer from revealing information relating to the representation of a client. This prohibition also applies to disclosures by a lawyer that do not in themselves reveal protected information but could reasonably lead to the discovery of such information by a third person. A lawyer's use of a hypothetical to discuss issues relating to the representation is permissible so long as there is no reasonable likelihood that the listener will be able to ascertain the identity of the client or the situation involved.

**Authorized Disclosure**

[5] Except to the extent that the client's instructions or special circumstances limit that authority, a lawyer is impliedly authorized to make disclosures about a client when appropriate in carrying out the representation. In some situations, for example, a lawyer may be impliedly authorized to admit a fact that cannot properly be disputed or to make a disclosure that facilitates a satisfactory conclusion to a matter. Lawyers in a firm may, in the course of the firm's practice, disclose to each other information relating to a client of the firm, unless the client has instructed that particular information be confined to specified lawyers.

**Disclosure Adverse to Client**

[6] Although the public interest is usually best served by a strict rule requiring lawyers to preserve the confidentiality of information relating to the representation of their clients, the confidentiality rule is subject to limited exceptions. Paragraph (b)(1) recognizes the overriding value of life and physical integrity and permits disclosure reasonably necessary to prevent reasonably certain death or substantial bodily harm. Such harm is reasonably certain to occur if it will be suffered imminently or if there is a present and substantial threat that a person will suffer such harm at a later date if the lawyer fails to take action necessary to eliminate the threat. Thus, a lawyer who knows that a client has accidentally discharged toxic waste into a town's water supply may reveal this information to the authorities if there is a present and substantial risk that a person who drinks the water will contract a life-threatening or debilitating
disease and the lawyer's disclosure is necessary to eliminate the threat or reduce the number of victims.

[7] Paragraph (b)(2) is a limited exception to the rule of confidentiality that permits the lawyer to reveal information to the extent necessary to enable affected persons or appropriate authorities to prevent the client from committing a crime or fraud, as defined in Rule 1.0(d), that is reasonably certain to result in substantial injury to the financial or property interests of another and in furtherance of which the client has used or is using the lawyer’s services. Such a serious abuse of the client-lawyer relationship by the client forfeits the protection of this Rule. The client can, of course, prevent such disclosure by refraining from the wrongful conduct. Although paragraph (b)(2) does not require the lawyer to reveal the client’s misconduct, the lawyer may not counsel or assist the client in conduct the lawyer knows is criminal or fraudulent. See Rule 1.2(d). See also Rule 1.16 with respect to the lawyer’s obligation or right to withdraw from the representation of the client in such circumstances, and Rule 1.13(c), which permits the lawyer, where the client is an organization, to reveal information relating to the representation in limited circumstances.

[8] Paragraph (b)(3) addresses the situation in which the lawyer does not learn of the client’s crime or fraud until after it has been consummated. Although the client no longer has the option of preventing disclosure by refraining from the wrongful conduct, there will be situations in which the loss suffered by the affected person can be prevented, rectified or mitigated. In such situations, the lawyer may disclose information relating to the representation to the extent necessary to enable the affected persons to prevent or mitigate reasonably certain losses or to attempt to recoup their losses. Paragraph (b)(3) does not apply when a person who has committed a crime or fraud thereafter employs a lawyer for representation concerning that offense.

[9] A lawyer's confidentiality obligations do not preclude a lawyer from securing confidential legal advice about the lawyer's personal responsibility to comply with these Rules. In most situations, disclosing information to secure such advice will be impliedly authorized for the lawyer to carry out the representation. Even when the disclosure is not impliedly authorized, paragraph (b)(4) permits such disclosure because of the importance of a lawyer’s compliance with the Rules of Professional Conduct.

[10] Where a legal claim or disciplinary charge alleges complicity of the lawyer in a client’s conduct or other misconduct of the lawyer involving representation of the client, the lawyer may respond to the extent the lawyer reasonably believes necessary to establish a defense. The same is true with respect to a claim involving the conduct or representation of a former client. Such a charge can arise in a civil, criminal, disciplinary or other proceeding and can be based on a wrong allegedly committed by the lawyer against the client or on a wrong alleged by a third person, for example, a person claiming to have been defrauded by the lawyer and client acting together. The lawyer’s right to respond arises when an assertion of such complicity has been made. Paragraph (b)(5) does not require the lawyer to await the commencement of an action or proceeding that charges such complicity, so that the defense may be established by responding directly to a third party who has made such an assertion. The right to defend also applies, of course, where a proceeding has been commenced.

[11] A lawyer entitled to a fee is permitted by paragraph (b)(5) to prove the services rendered in an action to collect it. This aspect of the rule expresses the principle that the beneficiary of a fiduciary relationship may not exploit it to the detriment of the fiduciary.

[12] Other law may require that a lawyer disclose information about a client. Whether such a law supersedes Rule 1.6 is a question of law beyond the scope of these Rules. When disclosure of information relating to the representation appears to be required by other law, the lawyer must discuss the matter with the client to the extent required by Rule 1.4. If, however, the other law supersedes this Rule and requires disclosure, paragraph (b)(6) permits the lawyer to make such disclosures as are necessary to comply with the law.
Detection of Conflicts of Interest

[13] Paragraph (b)(7) recognizes that lawyers in different firms may need to disclose limited information to each other to detect and resolve conflicts of interest, such as when a lawyer is considering an association with another firm, two or more firms are considering a merger, or a lawyer is considering the purchase of a law practice. See Rule 1.17, Comment [7]. Under these circumstances, lawyers and law firms are permitted to disclose limited information, but only once substantive discussions regarding the new relationship have occurred. Any such disclosure should ordinarily include no more than the identity of the persons and entities involved in a matter, a brief summary of the general issues involved, and information about whether the matter has terminated. Even this limited information, however, should be disclosed only to the extent reasonably necessary to detect and resolve conflicts of interest that might arise from the possible new relationship. Moreover, the disclosure of any information is prohibited if it would compromise the attorney-client privilege or otherwise prejudice the client (e.g., the fact that a corporate client is seeking advice on a corporate takeover that has not been publicly announced; that a person has consulted a lawyer about the possibility of divorce before the person's intentions are known to the person's spouse; or that a person has consulted a lawyer about a criminal investigation that has not led to a public charge). Under those circumstances, paragraph (a) prohibits disclosure unless the client or former client gives informed consent. A lawyer’s fiduciary duty to the lawyer’s firm may also govern a lawyer’s conduct when exploring an association with another firm and is beyond the scope of these Rules.

[14] Any information disclosed pursuant to paragraph (b)(7) may be used or further disclosed only to the extent necessary to detect and resolve conflicts of interest. Paragraph (b)(7) does not restrict the use of information acquired by means independent of any disclosure pursuant to paragraph (b)(7). Paragraph (b)(7) also does not affect the disclosure of information within a law firm when the disclosure is otherwise authorized, see Comment [5], such as when a lawyer in a firm discloses information to another lawyer in the same firm to detect and resolve conflicts of interest that could arise in connection with undertaking a new representation.

[15] A lawyer may be ordered to reveal information relating to the representation of a client by a court or by another tribunal or governmental entity claiming authority pursuant to other law to compel the disclosure. Absent informed consent of the client to do otherwise, the lawyer should assert on behalf of the client all nonfrivolous claims that the order is not authorized by other law or that the information sought is protected against disclosure by the attorney-client privilege or other applicable law. In the event of an adverse ruling, the lawyer must consult with the client about the possibility of appeal to the extent required by Rule 1.4. Unless review is sought, however, paragraph (b)(6) permits the lawyer to comply with the court's order.

[16] Paragraph (b) permits disclosure only to the extent the lawyer reasonably believes the disclosure is necessary to accomplish one of the purposes specified. Where practicable, the lawyer should first seek to persuade the client to take suitable action to obviate the need for disclosure. In any case, a disclosure adverse to the client's interest should be no greater than the lawyer reasonably believes necessary to accomplish the purpose. If the disclosure will be made in connection with a judicial proceeding, the disclosure should be made in a manner that limits access to the information to the tribunal or other persons having a need to know it and appropriate protective orders or other arrangements should be sought by the lawyer to the fullest extent practicable.

[17] Paragraph (b) permits but does not require the disclosure of information relating to a client's representation to accomplish the purposes specified in paragraphs (b)(1) through (b)(6). In exercising the discretion conferred by this Rule, the lawyer may consider such factors as the nature of the lawyer's relationship with the client and with those who might be injured by
the client, the lawyer's own involvement in the transaction and factors that may extenuate the conduct in question. A lawyer's decision not to disclose as permitted by paragraph (b) does not violate this Rule. Disclosure may be required, however, by other Rules. Some Rules require disclosure only if such disclosure would be permitted by paragraph (b). See Rules 1.2(d), 4.1(b), 8.1 and 8.3. Rule 3.3, on the other hand, requires disclosure in some circumstances regardless of whether such disclosure is permitted by this Rule. See Rule 3.3(c).

Acting Competently to Preserve Confidentiality

[18] Paragraph (c) requires a lawyer to act competently to safeguard information relating to the representation of a client against unauthorized access by third parties and against inadvertent or unauthorized disclosure by the lawyer or other persons who are participating in the representation of the client or who are subject to the lawyer’s supervision. See Rules 1.1, 5.1 and 5.3. The unauthorized access to, or the inadvertent or unauthorized disclosure of, information relating to the representation of a client does not constitute a violation of paragraph (c) if the lawyer has made reasonable efforts to prevent the access or disclosure. Factors to be considered in determining the reasonableness of the lawyer’s efforts include, but are not limited to, the sensitivity of the information, the likelihood of disclosure if additional safeguards are not employed, the cost of employing additional safeguards, the difficulty of implementing the safeguards, and the extent to which the safeguards adversely affect the lawyer’s ability to represent clients (e.g., by making a device or important piece of software excessively difficult to use). A client may require the lawyer to implement special security measures not required by this Rule or may give informed consent to forgo security measures that would otherwise be required by this Rule. Whether a lawyer may be required to take additional steps to safeguard a client’s information in order to comply with other law, such as state and federal laws that govern data privacy or that impose notification requirements upon the loss of, or unauthorized access to, electronic information, is beyond the scope of these Rules. For a lawyer’s duties when sharing information with nonlawyers outside the lawyer’s own firm, see Rule 5.3, Comments [3]-[4].

[19] When transmitting a communication that includes information relating to the representation of a client, the lawyer must take reasonable precautions to prevent the information from coming into the hands of unintended recipients. This duty, however, does not require that the lawyer use special security measures if the method of communication affords a reasonable expectation of privacy. Special circumstances, however, may warrant special precautions. Factors to be considered in determining the reasonableness of the lawyer's expectation of confidentiality include the sensitivity of the information and the extent to which the privacy of the communication is protected by law or by a confidentiality agreement. A client may require the lawyer to implement special security measures not required by this Rule or may give informed consent to the use of a means of communication that would otherwise be prohibited by this Rule. Whether a lawyer may be required to take additional steps in order to comply with other law, such as state and federal laws that govern data privacy, is beyond the scope of these Rules.

Former Client

[20] The duty of confidentiality continues after the client-lawyer relationship has terminated. See Rule 1.9(c)(2). See Rule 1.9(c)(1) for the prohibition against using such information to the disadvantage of the former client.

Definitional Cross-References
“Firm” See Rule 1.0(c)
“Fraud” See Rule 1.0(d)
“Informed consent” See Rule 1.0(e)
“Reasonably” See Rule 1.0(h)
“Reasonably believes” See Rule 1.0(i)
“Substantial” See Rule 1.0(l)
RULE 1.8: CONFLICT OF INTEREST: CURRENT CLIENTS: SPECIFIC RULES

(a) A lawyer shall not enter into a business transaction with a client or knowingly acquire an ownership, possessory, security or other pecuniary interest adverse to a client unless:

(1) the transaction and terms on which the lawyer acquires the interest are fair and reasonable to the client and are fully disclosed and transmitted in writing in a manner that can be reasonably understood by the client;

(2) the client is advised in writing of the desirability of seeking and is given a reasonable opportunity to seek the advice of independent legal counsel on the transaction; and

(3) the client gives informed consent, in a writing signed by the client, to the essential terms of the transaction and the lawyer's role in the transaction, including whether the lawyer is representing the client in the transaction.

(b) A lawyer shall not use information relating to representation of a client to the disadvantage of the client unless the client gives informed consent, except as permitted or required by these Rules.

(c) A lawyer shall not solicit any substantial gift from a client, including a testamentary gift, or prepare on behalf of a client an instrument giving the lawyer or a person related to the lawyer any substantial gift unless the lawyer or other recipient of the gift is related to the client. For purposes of this paragraph, related persons include a spouse, child, grandchild, parent, grandparent or other relative or individual with whom the lawyer or the client maintains a close, familial relationship.

(d) Prior to the conclusion of representation of a client, a lawyer shall not make or negotiate an agreement giving the lawyer literary or media rights to a portrayal or account based in substantial part on information relating to the representation.

(e) A lawyer shall not provide financial assistance to a client in connection with pending or contemplated litigation, except that:

(1) a lawyer may advance court costs and expenses of litigation, the repayment of which may be contingent on the outcome of the matter; and

(2) a lawyer representing an indigent client may pay court costs and expenses of litigation on behalf of the client.

(f) A lawyer shall not accept compensation for representing a client from one other than the client unless:

(1) the client gives informed consent;

(2) there is no interference with the lawyer's independence of professional judgment or with the client-lawyer relationship; and

(3) information relating to representation of a client is protected as
required by Rule 1.6.

(g) A lawyer who represents two or more clients shall not participate in making an aggregate settlement of the claims of or against the clients, or in a criminal case an aggregated agreement as to guilty or nolo contendere pleas, unless each client gives informed consent, in a writing signed by the client. The lawyer's disclosure shall include the existence and nature of all the claims or pleas involved and of the participation of each person in the settlement.

(h) A lawyer shall not:

(1) make an agreement prospectively limiting the lawyer's liability to a client for malpractice unless the client is independently represented in making the agreement; or

(2) settle a claim or potential claim for such liability with an unrepresented client or former client unless that person is advised in writing of the desirability of seeking and is given a reasonable opportunity to seek the advice of independent legal counsel in connection therewith.

(i) A lawyer shall not acquire a proprietary interest in the cause of action or subject matter of litigation the lawyer is conducting for a client, except that the lawyer may:

(1) acquire a lien authorized by law to secure the lawyer's fee or expenses; and

(2) contract with a client for a reasonable contingent fee in a civil case.

(j) A lawyer shall not have sexual relations with a client unless a consensual sexual relationship existed between them when the client-lawyer relationship commenced.

(k) While lawyers are associated in a firm, a prohibition in the foregoing paragraphs (a) through (i) that applies to any one of them shall apply to all of them.

Comment

Business Transactions between Client and Lawyer

[1] A lawyer's legal skill and training, together with the relationship of trust and confidence between lawyer and client, create the possibility of overreaching when the lawyer participates in a business, property or financial transaction with a client, for example, a loan or sales transaction or a lawyer investment on behalf of a client. The requirements of paragraph (a) must be met even when the transaction is not closely related to the subject matter of the representation, as when a lawyer drafting a will for a client learns that the client needs money for unrelated expenses and offers to make a loan to the client. The Rule applies to lawyers engaged in the sale of goods or services related to the practice of law, for example, the sale of title insurance or investment services to existing clients of the lawyer's legal practice. See Rule 5.7. It also applies to lawyers purchasing property from estates they represent. It does not apply to ordinary fee arrangements between client and lawyer, which are governed by Rule 1.5, although its requirements must be met when the lawyer accepts an interest in the client's business or other
nonmonetary property as payment of all or part of a fee. In addition, the Rule does not apply to
standard commercial transactions between the lawyer and the client for products or services that
the client generally markets to others, for example, banking or brokerage services, medical
services, products manufactured or distributed by the client, and utilities' services. In such
transactions, the lawyer has no advantage in dealing with the client, and the restrictions in
paragraph (a) are unnecessary and impracticable.

[2] Paragraph (a)(1) requires that the transaction itself be fair to the client and that its
esential terms be communicated to the client, in writing, in a manner that can be reasonably
understood. Paragraph (a)(2) requires that the client also be advised, in writing, of the desirability
of seeking the advice of independent legal counsel. It also requires that the client be given a
reasonable opportunity to obtain such advice. Paragraph (a)(3) requires that the lawyer obtain
the client's informed consent, in a writing signed by the client, both to the essential terms of the
transaction and to the lawyer's role. When necessary, the lawyer should discuss both the
material risks of the proposed transaction, including any risk presented by the lawyer's
involvement, and the existence of reasonably available alternatives and should explain why the
advice of independent legal counsel is desirable. See Rule 1.0(e) (definition of informed
consent).

[3] The risk to a client is greatest when the client expects the lawyer to represent
the client in the transaction itself or when the lawyer's financial interest otherwise poses a significant
risk that the lawyer's representation of the client will be materially limited by the lawyer's
financial interest in the transaction. Here the lawyer's role requires that the lawyer must comply,
not only with the requirements of paragraph (a), but also with the requirements of Rule 1.7.
Under that Rule, the lawyer must disclose the risks associated with the lawyer's dual role as both
legal adviser and participant in the transaction, such as the risk that the lawyer will structure the
transaction or give legal advice in a way that favors the lawyer's interests at the expense of the
client. Moreover, the lawyer must obtain the client's informed consent. In some cases, the
lawyer's interest may be such that Rule 1.7 will preclude the lawyer from seeking the client's
consent to the transaction.

[4] If the client is independently represented in the transaction, paragraph (a)(2) of
this Rule is inapplicable, and the paragraph (a)(1) requirement for full disclosure is satisfied
either by a written disclosure by the lawyer involved in the transaction or by the client's
independent counsel. The fact that the client was independently represented in the transaction is
relevant in determining whether the agreement was fair and reasonable to the client as paragraph
(a)(1) further requires.

Use of Information Related to Representation

[5] Use of information relating to the representation to the disadvantage of the client
violates the lawyer's duty of loyalty. Paragraph (b) applies when the information is used to
benefit either the lawyer or a third person, such as another client or business associate of the
lawyer. For example, if a lawyer learns that a client intends to purchase and develop several
parcels of land, the lawyer may not use that information to purchase one of the parcels in
competition with the client or to recommend that another client make such a purchase. The Rule
does not prohibit uses that do not disadvantage the client. For example, a lawyer who learns a
government agency's interpretation of trade legislation during the representation of one client
may properly use that information to benefit other clients. Paragraph (b) prohibits
disadvantageous use of client information unless the client gives informed consent, except as
permitted or required by these Rules. See Rules 1.2(d), 1.6, 1.9(c), 3.3, 4.1(b), 8.1 and 8.3.

Gifts to Lawyers
[6] A lawyer may accept a gift from a client, if the transaction meets general standards of fairness. For example, a simple gift such as a present given at a holiday or as a token of appreciation is permitted. If a client offers the lawyer a more substantial gift, paragraph (c) does not prohibit the lawyer from accepting it, although such a gift may be voidable by the client under the doctrine of undue influence, which treats client gifts as presumptively fraudulent. In any event, due to concerns about overreaching and imposition on clients, a lawyer may not suggest that a substantial gift be made to the lawyer or for the lawyer's benefit, except where the lawyer is related to the client as set forth in paragraph (c).

[7] If effectuation of a substantial gift requires preparing a legal instrument such as a will or conveyance, the client should have the detached advice that another lawyer can provide. The sole exception to this Rule is where the client is a relative of the donee.

[8] This Rule does not prohibit a lawyer from seeking to have the lawyer or a partner or associate of the lawyer named as executor of the client's estate or to another potentially lucrative fiduciary position. Nevertheless, such appointments will be subject to the general conflict of interest provision in Rule 1.7 when there is a significant risk that the lawyer's interest in obtaining the appointment will materially limit the lawyer's independent professional judgment in advising the client concerning the choice of an executor or other fiduciary. In obtaining the client's informed consent to the conflict, the lawyer should advise the client concerning the nature and extent of the lawyer's financial interest in the appointment, as well as the availability of alternative candidates for the position.

**Literary Rights**

[9] An agreement by which a lawyer acquires literary or media rights concerning the conduct of the representation creates a conflict between the interests of the client and the personal interests of the lawyer. Measures suitable in the representation of the client may detract from the publication value of an account of the representation. Paragraph (d) does not prohibit a lawyer representing a client in a transaction concerning literary property from agreeing that the lawyer's fee shall consist of a share in ownership in the property, if the arrangement conforms to Rule 1.5 and paragraphs (a) and (i).

**Financial Assistance**

[10] Lawyers may not subsidize lawsuits or administrative proceedings brought on behalf of their clients, including making or guaranteeing loans to their clients for living expenses, because to do so would encourage clients to pursue lawsuits that might not otherwise be brought and because such assistance gives lawyers too great a financial stake in the litigation. These dangers do not warrant a prohibition on a lawyer lending a client court costs and litigation expenses, including the expenses of medical examination and the costs of obtaining and presenting evidence, because these advances are virtually indistinguishable from contingent fees and help ensure access to the courts. Similarly, an exception allowing lawyers representing indigent clients to pay court costs and litigation expenses regardless of whether these funds will be repaid is warranted.

**Person Paying for a Lawyer's Services**
[11] Lawyers are frequently asked to represent a client under circumstances in which a third person will compensate the lawyer, in whole or in part. The third person might be a relative or friend, an indemnitor (such as a liability insurance company) or a co-client (such as a corporation sued along with one or more of its employees). Because third-party payers frequently have interests that differ from those of the client, including interests in minimizing the amount spent on the representation and in learning how the representation is progressing, lawyers are prohibited from accepting or continuing such representations unless the lawyer determines that there will be no interference with the lawyer's independent professional judgment and there is informed consent from the client. See also Rule 5.4(c) (prohibiting interference with a lawyer's professional judgment by one who recommends, employs or pays the lawyer to render legal services for another).

[12] Sometimes, it will be sufficient for the lawyer to obtain the client's informed consent regarding the fact of the payment and the identity of the third-party payer. If, however, the fee arrangement creates a conflict of interest for the lawyer, then the lawyer must comply with Rule 1.7. The lawyer must also conform to the requirements of Rule 1.6 concerning confidentiality. Under Rule 1.7(a), a conflict of interest exists if there is significant risk that the lawyer's representation of the client will be materially limited by the lawyer's own interest in the fee arrangement or by the lawyer's responsibilities to the third-party payer (for example, when the third-party payer is a co-client). Under Rule 1.7(b), the lawyer may accept or continue the representation with the informed consent of each affected client, unless the conflict is nonconsentable under that paragraph. Under Rule 1.7(b), the informed consent must be confirmed in writing.

Aggregate Settlements

[13] Differences in willingness to make or accept an offer of settlement are among the risks of common representation of multiple clients by a single lawyer. Under Rule 1.7, this is one of the risks that should be discussed before undertaking the representation, as part of the process of obtaining the clients' informed consent. In addition, Rule 1.2(a) protects each client's right to have the final say in deciding whether to accept or reject an offer of settlement and in deciding whether to enter a guilty or nolo contendere plea in a criminal case. The rule stated in this paragraph is a corollary of both these Rules and provides that, before any settlement offer or plea bargain is made or accepted on behalf of multiple clients, the lawyer must inform each of them about all the material terms of the settlement, including what the other clients will receive or pay if the settlement or plea offer is accepted. See also Rule 1.0(e) (definition of informed consent). Lawyers representing a class of plaintiffs or defendants, or those proceeding derivatively, may not have a full client-lawyer relationship with each member of the class; nevertheless, such lawyers must comply with applicable rules regulating notification of class members and other procedural requirements designed to ensure adequate protection of the entire class.

Limiting Liability and Settling Malpractice Claims

[14] Agreements prospectively limiting a lawyer's liability for malpractice are prohibited unless the client is independently represented in making the agreement because they are likely to undermine competent and diligent representation. Also, many clients are unable to evaluate the desirability of making such an agreement before a dispute has arisen, particularly if
they are then represented by the lawyer seeking the agreement. This paragraph does not, however, prohibit a lawyer from entering into an agreement with the client to arbitrate legal malpractice claims, provided such agreements are enforceable and the client is fully informed of the scope and effect of the agreement. Nor does this paragraph limit the ability of lawyers to practice in the form of a limited-liability entity, where permitted by law, provided that each lawyer remains personally liable to the client for his or her own conduct and the firm complies with any conditions required by law, such as provisions requiring client notification or maintenance of adequate liability insurance. Nor does it prohibit an agreement in accordance with Rule 1.2 that defines the scope of the representation, although a definition of scope that makes the obligations of representation illusory will amount to an attempt to limit liability.

[15] Agreements settling a claim or a potential claim for malpractice are not prohibited by this Rule. Nevertheless, in view of the danger that a lawyer will take unfair advantage of an unrepresented client or former client, the lawyer must first advise such a person in writing of the appropriateness of independent representation in connection with such a settlement. In addition, the lawyer must give the client or former client a reasonable opportunity to find and consult independent counsel.

Acquiring Proprietary Interest in Litigation

[16] Paragraph (i) states the traditional general rule that lawyers are prohibited from acquiring a proprietary interest in litigation. Like paragraph (e), the general rule has its basis in common law champerty and maintenance and is designed to avoid giving the lawyer too great an interest in the representation. In addition, when the lawyer acquires an ownership interest in the subject of the representation, it will be more difficult for a client to discharge the lawyer if the client so desires. The Rule is subject to specific exceptions developed in decisional law and continued in these Rules. The exception for certain advances of the costs of litigation is set forth in paragraph (e). In addition, paragraph (i) sets forth exceptions for liens authorized by law to secure the lawyer's fees or expenses and contracts for reasonable contingent fees. The law of each jurisdiction determines which liens are authorized by law. These may include liens granted by statute, liens originating in common law and liens acquired by contract with the client. When a lawyer acquires by contract a security interest in property other than that recovered through the lawyer's efforts in the litigation, such an acquisition is a business or financial transaction with a client and is governed by the requirements of paragraph (a). Contracts for contingent fees in civil cases are governed by Rule 1.5.

Client-Lawyer Sexual Relationships

[17] The relationship between lawyer and client is a fiduciary one in which the lawyer occupies the highest position of trust and confidence. The relationship is almost always unequal; thus, a sexual relationship between lawyer and client can involve unfair exploitation of the lawyer's fiduciary role, in violation of the lawyer's basic ethical obligation not to use the trust of the client to the client's disadvantage. In addition, such a relationship presents a significant danger that, because of the lawyer's emotional involvement, the lawyer will be unable to represent the client without impairment of the exercise of independent professional judgment. Moreover, a blurred line between the professional and personal relationships may make it difficult to predict to what extent client confidences will be protected by the attorney-client
evidentiary privilege, since client confidences are protected by privilege only when they are imparted in the context of the client-lawyer relationship. Because of the significant danger of harm to client interests and because the client's own emotional involvement renders it unlikely that the client could give adequate informed consent, this Rule prohibits the lawyer from having sexual relations with a client regardless of whether the relationship is consensual and regardless of the absence of prejudice to the client.

[18] Sexual relationships that predate the client-lawyer relationship are not prohibited. Issues relating to the exploitation of the fiduciary relationship and client dependency are diminished when the sexual relationship existed prior to the commencement of the client-lawyer relationship. However, before proceeding with the representation in these circumstances, the lawyer should consider whether the lawyer's ability to represent the client will be materially limited by the relationship. See Rule 1.7(a)(2).

[19] When the client is an organization, paragraph (j) of this Rule prohibits a lawyer for the organization (whether inside counsel or outside counsel) from having a sexual relationship with a constituent of the organization who supervises, directs or regularly consults with that lawyer concerning the organization's legal matters.

**Imputation of Prohibitions**

[20] Under paragraph (k), a prohibition on conduct by an individual lawyer in paragraphs (a) through (i) also applies to all lawyers associated in a firm with the personally prohibited lawyer. For example, one lawyer in a firm may not enter into a business transaction with a client of another member of the firm without complying with paragraph (a), even if the first lawyer is not personally involved in the representation of the client. The prohibition set forth in paragraph (j) is personal and is not applied to associated lawyers.

**Definitional Cross-References**

“Firm” See Rule 1.0(c)
“Informed consent” See Rule 1.0(e)
“Knowingly” See Rule 1.0(f)
“Substantial” See Rule 1.0(l)
“Writing” and “Signed” See Rule 1.0(n)
RULE 1.9: DUTIES TO FORMER CLIENTS

(a) A lawyer who has formerly represented a client in a matter shall not thereafter represent another person in the same or a substantially related matter in which that person’s interests are materially adverse to the interests of the former client unless the former client gives informed consent, confirmed in writing.

(b) A lawyer shall not knowingly represent a person in the same or a substantially related matter in which a firm with which the lawyer formerly was associated had previously represented a client

   (1) whose interests are materially adverse to that person; and

   (2) about whom the lawyer had acquired information protected by Rules 1.6 and 1.9(c) that is material to the matter;

unless the former client gives informed consent, confirmed in writing.

(c) A lawyer who has formerly represented a client in a matter or whose present or former firm has formerly represented a client shall not thereafter:

   (1) use information relating to the representation to the disadvantage of the former client except as these Rules would permit or require with respect to a client, or when the information has become generally known; or

   (2) reveal information relating to the representation except as these Rules would permit or require with respect to a client.

Comment

[1] After termination of a client-lawyer relationship, a lawyer has certain continuing duties with respect to confidentiality and conflicts of interest and thus may not represent another client except in conformity with this Rule. Under this Rule, for example, a lawyer could not properly seek to rescind on behalf of a new client a contract drafted on behalf of the former client. So also a lawyer who has prosecuted an accused person could not properly represent the accused in a subsequent civil action against the government concerning the same transaction. Nor could a lawyer who has represented multiple clients in a matter represent one of the clients against the others in the same or a substantially related matter after a dispute arose among the clients in that matter, unless all affected clients give informed consent. See Comment [9]. Current and former government lawyers must comply with this Rule to the extent required by Rule 1.11.

[2] The scope of a "matter" for purposes of this Rule depends on the facts of a particular situation or transaction. The lawyer's involvement in a matter can also be a question of degree. When a lawyer has been directly involved in a specific transaction, subsequent representation of other clients with materially adverse interests in that transaction clearly is prohibited. On the other hand, a lawyer who recurrently handled a type of problem for a former client is not precluded from later representing another client in a factually distinct problem of that type even though the subsequent representation involves a position adverse to the prior client. Similar considerations can apply to the reassignment of military lawyers between defense and prosecution functions within the same military jurisdictions. The underlying question is whether the lawyer was so involved in the matter that the subsequent representation can be justly
regarded as a changing of sides in the matter in question.

[3] Matters are "substantially related" for purposes of this Rule if they involve the same transaction or legal dispute or if there otherwise is a substantial risk that confidential factual information as would normally have been obtained in the prior representation would materially advance the client's position in the subsequent matter. For example, a lawyer who has represented a businessperson and learned extensive private financial information about that person may not then represent that person's spouse in seeking a divorce. Similarly, a lawyer who has previously represented a client in securing environmental permits to build a shopping center would be precluded from representing neighbors seeking to oppose rezoning of the property on the basis of environmental considerations; however, the lawyer would not be precluded, on the grounds of substantial relationship, from defending a tenant of the completed shopping center in resisting eviction for nonpayment of rent. Information that has been disclosed to the public or to other parties adverse to the former client ordinarily will not be disqualifying. Information acquired in a prior representation may have been rendered obsolete by the passage of time, a circumstance that may be relevant in determining whether two representations are substantially related. In the case of an organizational client, general knowledge of the client’s policies and practices ordinarily will not preclude a subsequent representation; on the other hand, knowledge of specific facts gained in a prior representation that are relevant to the matter in question ordinarily will preclude such a representation. A former client is not required to reveal the confidential information learned by the lawyer in order to establish a substantial risk that the lawyer has confidential information to use in the subsequent matter. A conclusion about the possession of such information may be based on the nature of the services the lawyer provided the former client and information that would in ordinary practice be learned by a lawyer providing such services.

Lawyers Moving Between Firms

[4] When lawyers have been associated within a firm but then end their association, the question of whether a lawyer should undertake representation is more complicated. There are several competing considerations. First, the client previously represented by the former firm must be reasonably assured that the principle of loyalty to the client is not compromised. Second, the rule should not be so broadly cast as to preclude other persons from having reasonable choice of legal counsel. Third, the rule should not unreasonably hamper lawyers from forming new associations and taking on new clients after having left a previous association. In this connection, it should be recognized that today many lawyers practice in firms, that many lawyers to some degree limit their practice to one field or another, and that many move from one association to another several times in their careers. If the concept of imputation were applied with unqualified rigor, the result would be radical curtailment of the opportunity of lawyers to move from one practice setting to another and of the opportunity of clients to change counsel.

[5] Paragraph (b) operates to disqualify the lawyer only when the lawyer involved has actual knowledge of information protected by Rules 1.6 and 1.9(c). Thus, if a lawyer while with one firm acquired no knowledge or information relating to a particular client of the firm, and that lawyer later joined another firm, neither the lawyer individually nor the second firm is disqualified from representing another client in the same or a related matter even though the interests of the two clients conflict. See Rule 1.10(b) for the restrictions on a firm once a lawyer has terminated association with the firm.

[6] Application of paragraph (b) depends on a situation's particular facts, aided by inferences, deductions or working presumptions that reasonably may be made about the way in which lawyers work together. A lawyer may have general access to files of all clients of a law firm and may regularly participate in discussions of their affairs; it should be inferred that such a lawyer in fact is privy to all information about all the firm's clients. In contrast, another lawyer
may have access to the files of only a limited number of clients and participate in discussions of the affairs of no other clients; in the absence of information to the contrary, it should be inferred that such a lawyer in fact is privy to information about the clients actually served but not those of other clients. In such an inquiry, the burden of proof should rest upon the firm whose disqualification is sought.

[7] Independent of the question of disqualification of a firm, a lawyer changing professional association has a continuing duty to preserve confidentiality of information about a client formerly represented. See Rules 1.6 and 1.9(c).

[8] Paragraph (c) provides that information acquired by the lawyer in the course of representing a client may not subsequently be used or revealed by the lawyer to the disadvantage of the client. However, the fact that a lawyer has once served a client does not preclude the lawyer from using generally known information about that client when later representing another client.

[9] The provisions of this Rule are for the protection of former clients and can be waived if the client gives informed consent, which consent must be confirmed in writing under paragraphs (a) and (b). See Rule 1.0(e). With regard to the effectiveness of an advance waiver, see Comment [22] to Rule 1.7. With regard to disqualification of a firm with which a lawyer is or was formerly associated, see Rule 1.10.

**Definitional Cross-References**

“Confirmed in writing” See Rule 1.0(b)

“Firm” See Rule 1.0(c)

“Informed Consent” See Rule 1.0(e)

“Knowingly” and “Known” See Rule 1.0(f)

“Writing” See Rule 1.0(n)
RULE 1.13: ORGANIZATION AS CLIENT

(a) A lawyer employed or retained by an organization represents the organization acting through its duly authorized constituents.

(b) If a lawyer for an organization knows that an officer, employee or other person associated with the organization is engaged in action, intends to act or refuses to act in a matter related to the representation that is a violation of a legal obligation to the organization, or a violation of law that reasonably might be imputed to the organization, and that is likely to result in substantial injury to the organization, then the lawyer shall proceed as is reasonably necessary in the best interest of the organization. Unless the lawyer reasonably believes that it is not necessary in the best interest of the organization to do so, the lawyer shall refer the matter to higher authority in the organization, including, if warranted by the circumstances, to the highest authority that can act on behalf of the organization as determined by applicable law.

(c) Except as provided in paragraph (d), if

(1) despite the lawyer's efforts in accordance with paragraph (b) the highest authority that can act on behalf of the organization insists upon or fails to address in a timely and appropriate manner an action or a refusal to act, that is clearly a violation of law; and

(2) the lawyer reasonably believes that the violation is reasonably certain to result in substantial injury to the organization,

then the lawyer may reveal information relating to the representation whether or not Rule 1.6 permits such disclosure, but only if and to the extent the lawyer reasonably believes necessary to prevent substantial injury to the organization.

(d) Paragraph (c) shall not apply with respect to information relating to a lawyer's representation of an organization to investigate an alleged violation of law, or to defend the organization or an officer, employee or other constituent associated with the organization against a claim arising out of an alleged violation of law.

(e) A lawyer who reasonably believes that he or she has been discharged because of the lawyer's actions taken pursuant to paragraphs (b) or (c), or who withdraws under circumstances that require or permit the lawyer to take action under either of those paragraphs, shall proceed as the lawyer reasonably believes necessary to assure that the organization's highest authority is informed of the lawyer's discharge or withdrawal.

(f) In dealing with an organization's directors, officers, employees, members, shareholders or other constituents, a lawyer shall explain the identity of the client when the lawyer knows or reasonably should know that the organization's interests are adverse to those of the constituents with whom the lawyer is dealing.

(g) A lawyer representing an organization may also represent any of its directors, officers, employees, members, shareholders or other constituents, subject to the provisions of Rule 1.7. If the organization's consent to the dual representation is required by Rule 1.7, the consent shall be given by an appropriate official of the organization other than the individual who is to be represented, or by the shareholders.
Comment

The Entity as the Client

[1] An organizational client is a legal entity, but it cannot act except through its officers, directors, employees, shareholders and other constituents. Officers, directors, employees and shareholders are the constituents of the corporate organizational client. The duties defined in this Comment apply equally to unincorporated associations. "Other constituents" as used in this Comment means the positions equivalent to officers, directors, employees and shareholders held by persons acting for organizational clients that are not corporations.

[2] When one of the constituents of an organizational client communicates with the organization's lawyer in that person's organizational capacity, the communication is protected by Rule 1.6. Thus, by way of example, if an organizational client requests its lawyer to investigate allegations of wrongdoing, interviews made in the course of that investigation between the lawyer and the client's employees or other constituents are covered by Rule 1.6. This does not mean, however, that constituents of an organizational client are the clients of the lawyer. The lawyer may not disclose to such constituents information relating to the representation except for disclosures explicitly or impliedly authorized by the organizational client in order to carry out the representation or as otherwise permitted by Rule 1.6.

[3] When constituents of the organization make decisions for it, the decisions ordinarily must be accepted by the lawyer even if their utility or prudence is doubtful. Decisions concerning policy and operations, including ones entailing serious risk, are not as such in the lawyer's province. Paragraph (b) makes clear, however, that when the lawyer knows that the organization is likely to be substantially injured by action of an officer or other constituent that violates a legal obligation to the organization or is in violation of law that might be imputed to the organization, the lawyer must proceed as is reasonably necessary in the best interest of the organization. As defined in Rule 1.0(f), knowledge can be inferred from circumstances, and a lawyer cannot ignore the obvious.

[4] In determining how to proceed under paragraph (b), the lawyer should give due consideration to the seriousness of the violation and its consequences, the responsibility in the organization and the apparent motivation of the person involved, the policies of the organization concerning such matters, and any other relevant considerations. Ordinarily, referral to a higher authority would be necessary. In some circumstances, however, it may be appropriate for the lawyer to ask the constituent to reconsider the matter; for example, if the circumstances involve a constituent’s innocent misunderstanding of law and subsequent acceptance of the lawyer’s advice, the lawyer may reasonably conclude that the best interest of the organization does not require that the matter be referred to higher authority. If a constituent persists in conduct contrary to the lawyer’s advice, it will be necessary for the lawyer to take steps to have the matter reviewed by a higher authority in the organization. If the matter is of sufficient seriousness and importance or urgency to the organization, referral to higher authority in the organization may be necessary even if the lawyer has not communicated with the constituent. Any measures taken should, to the extent practicable, minimize the risk of revealing information relating to the representation to persons outside the organization. Even in circumstances where a lawyer is not obligated by Rule 1.13 to proceed, a lawyer may bring to the attention of an organizational client, including its highest authority, matters that the lawyer reasonably believes to be of sufficient importance to warrant doing so in the best interest of the organization.

[5] Paragraph (b) also makes clear that when it is reasonably necessary to enable the organization to address the matter in a timely and appropriate manner, the lawyer must refer the matter to higher authority, including, if warranted by the circumstances, the highest authority that can act on behalf of the organization under applicable law. The organization's highest authority
to whom a matter may be referred ordinarily will be the board of directors or similar governing body. However, applicable law may prescribe that under certain conditions the highest authority reposes elsewhere, for example, in the independent directors of a corporation.

Relation to Other Rules

[6] The authority and responsibility provided in this Rule are concurrent with the authority and responsibility provided in other Rules. In particular, this Rule does not limit or expand the lawyer's responsibility under Rules 1.8, 1.16, 3.3 or 4.1. Paragraph (c) of this Rule supplements Rule 1.6(b) by providing an additional basis upon which the lawyer may reveal information relating to the representation, but does not modify, restrict, or limit the provisions of Rule 1.6(b)(1) – (6). Under paragraph (c) the lawyer may reveal such information only when the organization’s highest authority insists upon or fails to address threatened or ongoing action that is clearly a violation of law, and then only to the extent the lawyer reasonably believes necessary to prevent reasonably certain substantial injury to the organization. It is not necessary that the lawyer’s services be used in furtherance of the violation, but it is required that the matter be related to the lawyer’s representation of the organization. If the lawyer's services are being used by an organization to further a crime or fraud by the organization, Rules 1.6(b)(2) and 1.6(b)(3) may permit the lawyer to disclose confidential information. In such circumstances Rule 1.2(d) may also be applicable, in which event, withdrawal from the representation under Rule 1.16(a)(1) may be required.

[7] Paragraph (d) makes clear that the authority of a lawyer to disclose information relating to a representation in circumstances described in paragraph (c) does not apply with respect to information relating to a lawyer’s engagement by an organization to investigate an alleged violation of law or to defend the organization or an officer, employee or other person associated with the organization against a claim arising out of an alleged violation of law. This is necessary in order to enable organizational clients to enjoy the full benefits of legal counsel in conducting an investigation or defending against a claim.

[8] A lawyer who reasonably believes that he or she has been discharged because of the lawyer’s actions taken pursuant to paragraph (b) or (c), or who withdraws in circumstances that require or permit the lawyer to take action under either of these paragraphs, must proceed as the lawyer reasonably believes necessary to assure that the organization’s highest authority is informed of the lawyer’s discharge or withdrawal.

Government Agency

[9] The duty defined in this Rule applies to governmental organizations. Defining precisely the identity of the client and prescribing the resulting obligations of such lawyers may be more difficult in the government context and is a matter beyond the scope of these Rules. See Scope [18]. Although in some circumstances the client may be a specific agency, it may also be a branch of government, such as the executive branch, or the government as a whole. For example, if the action or failure to act involves the head of a bureau, either the department of which the bureau is a part or the relevant branch of government may be the client for purposes of this Rule. Moreover, in a matter involving the conduct of government officials, a government lawyer may have authority under applicable law to question such conduct more extensively than that of a lawyer for a private organization in similar circumstances. Thus, when the client is a governmental organization, a different balance may be appropriate between maintaining confidentiality and assuring that the wrongful act is prevented or rectified, for public business is involved. In addition, duties of lawyers employed by the government or lawyers in military service may be defined by statutes and regulation. This Rule does not limit that authority. See Scope.
Clarifying the Lawyer's Role

[10] There are times when the organization's interest may be or become adverse to those of one or more of its constituents. In such circumstances the lawyer should advise any constituent, whose interest the lawyer finds adverse to that of the organization of the conflict or potential conflict of interest, that the lawyer cannot represent such constituent, and that such person may wish to obtain independent representation. Care must be taken to assure that the individual understands that, when there is such adversity of interest, the lawyer for the organization cannot provide legal representation for that constituent individual, and that discussions between the lawyer for the organization and the individual may not be privileged.

[11] Whether such a warning should be given by the lawyer for the organization to any constituent individual may turn on the facts of each case.

Dual Representation

[12] Paragraph (g) recognizes that a lawyer for an organization may also represent a principal officer or major shareholder.

Derivative Actions

[13] Under generally prevailing law, the shareholders or members of a corporation may bring suit to compel the directors to perform their legal obligations in the supervision of the organization. Members of unincorporated associations have essentially the same right. Such an action may be brought nominally by the organization, but usually is, in fact, a legal controversy over management of the organization.

[14] The question can arise whether counsel for the organization may defend such an action. The proposition that the organization is the lawyer's client does not alone resolve the issue. Most derivative actions are a normal incident of an organization's affairs, to be defended by the organization's lawyer like any other suit. However, if the claim involves serious charges of wrongdoing by those in control of the organization, a conflict may arise between the lawyer's duty to the organization and the lawyer's relationship with the board. In those circumstances, Rule 1.7 governs who should represent the directors and the organization.

Definitional Cross-References

“Knows” See Rule 1.0(f)
“Reasonably” See Rule 1.0(h)
“Reasonably believes” See Rule 1.0(i)
“Reasonably should know” See Rule 1.0(j)
“Substantial” See Rule 1.0(l)
FINANCIAL REWARDS FOR WHISTLEBLOWING LAWYERS

KATHLEEN CLARK*
NANCY J. MOORE**

Abstract: The federal government increasingly relies on whistleblowers to ferret out fraud, awarding over $4 billion to whistleblowers under the False Claims Act (“FCA”) and the Dodd-Frank Wall Street Reform and Consumer Protection Act (“Dodd-Frank”). May lawyers ethically seek these whistleblower rewards? Several lawyers have tried unsuccessfully to serve as FCA whistleblowers. Additional lawyers may be seeking whistleblower rewards under Dodd-Frank, but the secrecy of the award process prevents us from knowing whether they have sought or received awards. This is the first Article to analyze in-depth the key questions for determining whether a lawyer may seek a federal whistleblower award: (1) When may a lawyer disclose a client’s confidential information? (2) When does a lawyer’s obligation of loyalty preclude seeking a personal benefit by disclosing a crime or fraud? (3) Do federal whistleblower laws preempt state ethics standards? (4) Which state’s ethics law applies when several states have significant contacts with the matter? These questions are enormously complex. Confidentiality exceptions differ widely among states. Lawyers are bound not just by conflict of interest rules, but also by the common-law duty not to profit from a client’s confidential information. While several federal courts have summarily rejected FCA preemption of state ethics standards, none of them confronted the fact that the FCA preempts state law fiduciary and contractual duties that would prevent nonlawyer insiders from serving as whistleblowers.

© 2015, Kathleen Clark & Nancy J. Moore. All rights reserved.
* Professor of Law, Washington University in St. Louis, kathleen@wustl.edu.
** Professor of Law and Nancy Barton Scholar, Boston University School of Law.

We are grateful for the helpful comments received when presenting an earlier draft of this Article to the BU Law Faculty Workshop Series, Legal Ethics Scholars’ Roundtable, Washington University Law Faculty Workshop, Sixth International Legal Ethics Conference, ABA National Conference on Professional Responsibility, National Institute on the Civil False Claims Act and Qui Tam Enforcement, and the Association of Professional Responsibility Lawyers. We would also like to thank Daniel Hurson, Jordan Thomas, and Jason Zuckerman for their comments on earlier drafts. Special thanks to Michael Harper and Kristin Collins for helping us understand the preemption issue and to Louise Teitz for educating us on aspects of the choice of law issue and for her helpful comments on several earlier drafts.
INTRODUCTION

The federal government increasingly relies on whistleblowers to assist in the enforcement of legal norms. This reliance is reflected not just in statutes promising protection for whistleblowers that experience retaliation, but also in other statutes providing large financial incentives for whistleblowers. The oldest of these statutes is the federal False Claims Act (“FCA”), 1 originally enacted in 1863 to enable whistleblowers (often organizational insiders) to file qui tam lawsuits in the name of the federal government against companies that have made false claims for payment from the gov-

ernment. These whistleblowers (“relators”) have a right to 10–30% of any resulting verdict or settlement, and have been awarded more than $4 billion in the years since Congress strengthened the statute in 1986.\(^2\) Based in part on the FCA’s track record, Congress recently expanded the availability of whistleblower financial incentives by enacting the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 (“Dodd-Frank”), which required the Securities and Exchange Commission (“SEC”) to give financial awards to whistleblowers who report securities violations to the SEC.\(^3\) If a whistleblower’s tip results in sanctions of greater than $1 million, the whistleblower can receive between 10–30% of the sanction amount.\(^4\) The SEC receives thousands of these tips every year, and has issued awards reaching into eight figures.

Lawyers for companies that do business with the government and for publicly traded companies have access to the kind of information that a whistleblower would need to file a qui tam FCA lawsuit or to file a whistleblower tip with the SEC. May lawyers—like other organizational insiders—take advantage of these financial incentives? Neither the FCA nor Dodd-Frank specifically addresses this question. As the government’s reliance on whistleblowers has expanded, it is increasingly important to identify when lawyers—like others—may take advantage of these whistleblower incentives.

A handful of lawyers have sued their former clients as qui tam relators under the FCA, although to date none have been successful.\(^5\) Among the obstacles confronting lawyer-relators are their obligations of confidentiality and loyalty under applicable state ethics rules; indeed, three of these lawsuits were dismissed based on findings that the lawyers had violated their ethical duties under state law. Apparently relying on aspects of these FCA cases, the SEC’s recently enacted Dodd-Frank whistleblower regulations exclude information learned in the course of a lawyer-client relationship unless a lawyer is permitted to disclose that information under either state confidentiality rules or the regulations that the SEC promulgated under the Sarbanes-Oxley Act (“SOX”) of 2002.\(^6\) But the SEC’s Dodd-Frank regulations do not address whether lawyers are eligible to receive a whistleblower incentives.

---

\(^2\) CIVIL DIV., U.S. DEP’T OF JUSTICE, FRAUD STATISTICS—OVERVIEW 2 (May 8, 2015), http://www.justice.gov/file/fcastatspdf/download [http://perma.cc/TX4V-L98C]. The government has recovered more than $30 billion through these qui tam suits during the same period. Id.
\(^5\) See infra notes 26–77 and accompanying text. Lawyers have also sued non-client third parties based on information obtained while representing a client. See infra notes 170–197.
award when their conduct violates their loyalty obligations under state conflict of interest rules or fiduciary law.

Should lawyers be permitted to receive financial rewards under the FCA and Dodd-Frank whistleblower programs? There are significant financial disincentives to engaging in whistleblowing. It can result not just in the end of a job, but the end of a career. Whistleblower awards can counteract these disincentives, for lawyers as well as for other insiders. Indeed, the SEC might argue that SOX expanded lawyers’ confidentiality exceptions and granted them additional discretion to make whistleblowing disclosures. Yet years after the legislation’s enactment, there is little evidence that lawyers have actually made disclosures to prevent, mitigate, or rectify client fraud. Lawyers—like others—may need whistleblower awards to counteract the financial disincentives for blowing the whistle.

But a client-lawyer relationship is, in some respects, different from other relationships. Lawyers can play a critical role in ensuring that clients understand and comply with the law. Some argue that this distinctive role means that we should not grant whistleblower awards to a company’s lawyers, particularly when lawyers who would seek such awards may violate duties of confidentiality or loyalty, even if we grant such awards to company employees who violate similar confidentiality or loyalty duties under state law.

Despite the obvious importance of such questions, it is not our purpose to engage in a normative analysis of federal whistleblower rewards to lawyers. Rather, we believe that before the normative question can be properly addressed, we need a more detailed understanding of the complex issues raised when lawyers seek federal whistleblower awards. Our descriptive agenda includes detailing the nuances of both confidentiality and loyalty obligations under state ethics laws, which vary significantly from state to state, particularly with respect to confidentiality exceptions. We also briefly discuss possible federal preemption of state ethics laws and the confounding choice of law issues raised in an era when lawyers perform their work in multiple jurisdictions, often far removed from their state of licensure.

Part I of this Article examines the relevant ethics law in light of the operation of the FCA’s unusual qui tam litigation procedures for whistleblowers who sue in the name of the government. Part II does the same with

---


respect to the Dodd-Frank statute and the SEC regulations for its whistleblower award program. Within each of these sections, we address how lawyers’ professional obligations of confidentiality and loyalty may affect their ability to qualify for financial awards. After describing the particulars of the FCA and Dodd-Frank whistleblower reward programs, we begin our ethics analysis with a brief discussion of the few FCA cases that have addressed the confidentiality and loyalty obligations of lawyer-relators. Although these cases address some of the relevant issues applicable under both the FCA and Dodd-Frank, they do not address or fully explore the wide range of ethical issues that we identify as arising under applicable ethics law. We analyze these issues first under the American Bar Association (“ABA”) Model Rules of Professional Conduct and then under the significant state variations, which exist primarily with respect to confidentiality. Within each category, we address lawyers’ obligations to both current and former clients, not only when the target of the lawyer’s disclosure is the client itself, but also when the target is a third party about whom the lawyer acquired information while representing a client.

With respect to the lawyer’s obligation of confidentiality, one of the issues we consider is whether it is ever “reasonably necessary” for a lawyer to actively seek a whistleblower reward in order to “prevent, mitigate or rectify” the substantial economic harm that may result from a client’s crime or fraud, especially when to do so requires the lawyer to file and actively litigate an FCA lawsuit against a current or former client. We also explore whether and under what circumstances whistleblower rewards are justified in states that permit disclosure solely to prevent future wrongdoing, given that the federal reward programs are based on establishing a company’s past wrongdoing. We conclude that, contrary to the apparent view of the courts in the existing FCA cases, it may be difficult for lawyer-whistleblowers to avoid violating state confidentiality rules, even in jurisdictions that permit disclosure to rectify past wrongdoing.

It may be difficult, but not impossible. Thus, we also consider, as did a federal district court in a recent FCA case, whether lawyers’ obligations of loyalty affect lawyers seeking whistleblower awards, even when confidentiality rules do not prohibit the requisite disclosure. For example, we consider whether a lawyer may continue to represent a client while seeking a whistleblower award, even on a matter unrelated to the lawyer’s ongoing work.

---

9 There are no reported cases involving lawyer-whistleblowers under Dodd-Frank. Although there are some differences in applying applicable ethics law to the two statutory programs, the issues are similar. As a result, the FCA cases are helpful in analyzing the ethical obligations of both FCA and Dodd-Frank whistleblowers.

10 Model Rules of Prof’l Conduct r. 1.6(d)(3) (AM. BAR ASS’N 2015).
We also explore whether former-client conflict rules or common-law fiduciary duties prevent a lawyer from seeking a whistleblower award, as well as whether and when lawyers are obligated to inform their clients that they have disclosed damaging information to governmental authorities. Here we also conclude that, although there are many open issues, ethics law presents substantial obstacles to lawyers acting in pursuit of their own interests, even when confidentiality rules permit them to disclose for other purposes. This law includes both conflict of interest rules for current clients, which are particularly salient for lawyers seeking to take advantage of the anonymity promised by Dodd-Frank, and a common-law fiduciary duty that prohibits lawyers (and other fiduciaries) from profiting from the use of confidential client information. This common-law duty applies to both current and former representation and also precludes lawyers from pursuing whistleblower awards against non-client third parties without the client’s consent, even when doing so will not harm the client.

Part III briefly addresses whether the federal whistleblower incentives under the FCA and Dodd-Frank preempt any aspects of state ethics laws regarding confidentiality and loyalty that would prevent a lawyer from participating in these whistleblower incentive programs. The few FCA decisions that have addressed the lawyer-relator issue agree that there has been no such preemption. But their discussion of the issue is minimal and ignores other FCA cases denying a defendant permission to assert a counterclaim against a nonlawyer-relator for breach of contract or breach of fiduciary duty when the assertion of such claims would undermine the federal government’s strong interest in encouraging whistleblowers to come forward. These nonlawyer-relator cases do not explicitly use preemption analysis, but the result nevertheless appears to be that at least some nonlawyer obligations under state law are being preempted by the FCA. We then address whether lawyers’ obligations under state standards might be treated differently under the FCA than the obligations of nonlawyers. As for Dodd-Frank, the SEC regulations expressly provide that more restrictive state confidentiality standards are preempted by the preexisting SOX lawyer whistleblower regulations. But the Dodd-Frank regulations do not mention lawyers’ loyalty obligations under state conflict of interest rules or fiduciary law. As a result, it is unclear whether and to what extent those obligations are impliedly preempted by the Dodd-Frank whistleblower bounty program.

Assuming that at least some state ethics rules are not preempted by the FCA or Dodd-Frank, Part IV introduces the difficult choice of law issues that may arise as a result of the considerable variation in state confidentiality rules. Both the FCA and the Dodd-Frank award programs involve national companies with multiple offices, as well as in-house lawyers who may not be licensed in the state where they advise the company. As a result, a
lawyer is unable to predict with certainty which state’s ethics rules govern. Because the SEC’s Dodd-Frank regulation apparently preempts state confidentiality rules that are stricter than the SEC’s own SOX regulation, and because loyalty provisions do not differ significantly from state to state, we focus our choice of law discussion on the difficult issues that arise when a federal court attempts to determine the ethical propriety of a lawyer-relator’s disclosure of confidential client information in bringing a qui tam lawsuit. Several FCA cases have briefly addressed choice of law issues in such a national setting, but we conclude that these decisions do not adequately confront the complexities of determining not only whose choice of law rule controls—the federal district court, the forum state, or some other state—but also whether a litigation or nonlitigation choice of law rule should apply. Although we do not thoroughly explore the choice of law issues raised here, we recommend that federal courts consider developing their own federal common-law choice of law rule for FCA lawsuits, perhaps incorporating existing approaches such as the ABA Model Rules’ nonlitigation choice of law provision or the Restatement (Second) of Conflict of Laws’ agency provisions.

This Article concludes with a summary of our findings. In addition, although we do not address the normative question of whether lawyers should be entitled to seek whistleblower rewards, we express concern about whether it is ever appropriate, as is provided under Dodd-Frank, for determinations of lawyer eligibility to be conducted in secret, in a process largely insulated from judicial review.

I. QUI TAM WHISTLEBLOWER AWARDS UNDER THE FALSE CLAIMS ACT

A. Primer on the FCA

The FCA enables almost anyone to file a lawsuit in the name of the United States to recover monies from someone who made false claims for payment from the government.11 In an FCA case, the relator files a complaint with the district court under seal and provides the U.S. Department of Justice with the complaint and a “written disclosure of substantially all material evidence and information the person possesses.”12 The defendant does not re-

---

11 See 31 U.S.C. § 3729; JOHN T. BOESE, CIVIL FALSE CLAIMS AND QUI TAM ACTIONS § 4.01[B] (4th ed. 2011) (stating that “virtually anyone can be a qui tam relator”). The statute excludes current or former members of the armed forces from serving as a relator if they are suing another member of the armed forces based on that other member’s service. 31 U.S.C. § 3730(c)(1)–(2) (2012). The statute also excludes members of Congress, the Judiciary, and “senior executive branch officials” from being named as defendants where the suit is “based on evidence or information known to the Government when the action was brought.” Id.
ceive the complaint until the court lifts the seal.13 This gives the government an opportunity to investigate the relator’s allegations and determine whether to participate in the relator’s FCA case (or even to file criminal charges).

Ultimately, the government has four options.14 It can: (1) ask the court to dismiss the relator’s case,15 (2) settle the case prior to formal intervention,16 (3) intervene and take over the conduct of the lawsuit,17 or (4) decline to intervene, allowing the relator to conduct the lawsuit.18 The government intervenes in only 27% of FCA cases,19 but intervened cases account for almost all (about 97%) of qui tam recoveries.20

Courts generally view FCA suits as sounding in fraud, and therefore impose on FCA complaints the heightened pleading requirements of Rule 9(b) of the Federal Rules of Civil Procedure,21 which requires a complaint alleging fraud to “state with particularity the circumstances constituting fraud or mistake.”22 This means that prior to civil discovery, the relator must generally have in hand evidence of the specific false claims for payment to the government, what they were for, and who made them. As a result, the FCA relators who can meet this requirement are generally individuals who had access to and retained copies of specific information about an organization’s false claims. In other words, most FCA relators are organizational insiders.

The FCA statute does not require relators to have entirely clean hands. It excludes relators who have been “convicted of criminal conduct arising

---

13 Id.
14 BOESE, supra note 11, § 4.05.
16 Id. § 3730(c)(2)(B).
17 Id. § 3730(b)(4)(A).
18 Id. § 3730(b)(4)(B). If the government declines to intervene, the relator controls the litigation but, at the government’s request, must serve the Justice Department with all filings, enabling the government to monitor the proceedings. Id. § 3730(c)(3). In most of the cases where the government has declined to intervene, relators seek voluntary dismissal of the case. See David Kwok, Evidence from the False Claims Act: Does Private Enforcement Attract Excessive Litigation?, 42 PUB. CONT. L.J. 225, 239–40 (2013) (noting the low success rate, 9%, of non-intervened cases).
20 See CIVIL DIV., U.S. DEP’T OF JUSTICE, supra note 2, at 2 (from 1987 until 2014, the government recovered more than $30 billion through qui tam lawsuits, but only $1 billion came from non-intervened cases).
21 See Kathleen M. Boozang, The New Relators: In-House Counsel and Compliance Officers, 6 J. HEALTH & LIFE SCI. L. 16, 18 (2012); see also REUBEN A. GUTTMAN & JACOB R. KIRKHAM, GRANT & EISENHOFER, P.A., FRONTLOADING THE CASE: THEME & THEORY IN FALSE CLAIMS AND FRAUD LITIGATION 3 (2012) (on file with authors) (“Although the FCA is not technically a fraud statute, courts have almost unanimously required parties to plead in compliance with Rule 9(b).”).
22 FED. R. CIV. P. 9(b).
from his or her role in” the FCA violation, but that is a relatively low bar. The statute thus implicitly recognizes that some of the individuals most likely to possess the information necessary for an FCA case may have been involved in the FCA violation. As one of the framers of the original statute recognized in 1863, the qui tam provisions “are based upon the idea of ‘setting a rogue to catch a rogue.’”

Even a “rogue” can be eligible for a whistleblower award.

B. Lawyers’ Confidentiality Obligations and the FCA

This section explores whether lawyers’ confidentiality obligations restrict their ability to serve as FCA relators. We first examine how courts have addressed this issue in FCA cases involving lawyer-relators, and then discuss the confidentiality standards and exceptions found in the ABA Model Rules of Professional Conduct and in state variations of those rules.

1. FCA Case Law Regarding Lawyer-Relators

Of the nearly ten-thousand qui tam FCA cases filed since 1986, we were unable to find any case in which a lawyer-relator sued a current client. We did, however, identify five cases in which a lawyer-relator sued a former client. In each of those cases, the lawyer alleged that he first expressed concern internally within the client company about the alleged FCA violation.

25 See id. This feature of the statute has implications for whether FCA defendants should be able to bring counterclaims against relators. See infra notes 302–351 and accompanying text.
26 See CIVIL DIV., U.S. DEP’T OF JUSTICE, supra note 2, at 2 (providing that 9960 qui tam cases were filed from Fiscal Years 1987 through 2014).
27 Four of these cases were filed under the federal False Claims Act. Vt. Agency of Nat. Res. v. United States ex rel. Stevens, 529 U.S. 765 (2000); United States ex rel. Fair Lab. Practices Assocs. v. Quest Diagnostics, Inc. (FLPA II), 734 F.3d 154 (2d Cir. 2013); United States ex rel. Repko v. Guthrie Clinic, P.C. (Repko I), 490 F. App’x 502 (3d Cir. 2012); United States ex rel. Doe v. X Corp. (Doe), 862 F. Supp. 1502 (E.D. Va. 1994). A fifth was filed under California’s False Claims Act. Bury v. Cmty. Hosps. of Cent. Cal., No. F036667, 2002 Cal. App. Unpub. LEXIS 1035 (May 8, 2002). In each of these cases, the lawyer-client relationship ended before the lawyer-relator filed his FCA lawsuit. Additional FCA cases have been brought by licensed lawyers who learned about the alleged fraud while working as compliance officers rather than as lawyers. See, e.g., United States ex rel. Frazier v. IASIS Healthcare Corp., No. 2:05-cv-766-RCJ, 2012 U.S. Dist. LEXIS 6896 (D. Az. Jan. 9, 2012). One legal commentator has asserted that such compliance officers are bound by lawyers’ professional confidentiality duties, see Boozang, supra note 21, at 10, but a recent bar ethics opinion concludes that they are not, see N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 14.
violation and that the client then retaliated against him.28 The government declined to intervene in any of these cases, and courts dismissed them before trial. Two of the cases were dismissed on grounds unrelated to legal ethics.29 In the remaining three, courts expressly evaluated how state confidentiality standards applied to the lawyer-relators, dismissing the cases because applicable state ethics rules prohibited the lawyer-relator from disclosing the information necessary to move forward with the FCA lawsuit.30 None of the courts ruled that lawyers were per se prohibited from serving as relators.31

The first FCA case in which a court applied lawyer confidentiality standards to a lawyer-relator was United States ex rel. Doe v. X Corp., in 1994.32 Lawyer-relator Doe worked in-house for a government contractor and alleged that the contractor violated the FCA by failing to disclose that the computers it sold contained remanufactured (rather than new) components.33 Before filing the FCA lawsuit, the lawyer raised these concerns internally, and the company disclosed additional information to the federal government.34 But after the company terminated the lawyer,35 he threatened

28 Under Seal v. Under Seal, 17 F.3d 1435 (4th Cir. 1994) (per curiam) (affirming the district court’s dismissal of Doe’s wrongful termination lawsuit because Virginia’s public policy exception to employment at will did not extend to lawyers); X Corp. v. Doe (X Corp. II), 816 F. Supp. 1086 (E.D. Va. 1993) (dismissing Doe’s 31 U.S.C. § 3730(h) retaliation claim), aff’d sub nom. Under Seal, 17 F.3d 1435. Before filing his FCA lawsuit, the lawyer-whistleblower in Stevens was fired by the Vermont Agency of Natural Resources after raising his concerns internally. Telephone Interview with Jonathan Stevens (Aug. 2, 2012); see also Stevens, 529 U.S. at 770 (noting that Stevens brought the action against his former employer). Bury filed a wrongful termination action against his former client. Bury, 2002 Cal. App. Unpub. LEXIS 1035, at *3. Repko’s retaliation claim under the FCA was dismissed on statute of limitations grounds. United States ex rel. Repko v. Guthrie Clinic, P.C., 557 F. Supp. 2d 522, 528–29 (M.D. Pa. 2008), aff’d, 490 F. App’x 502. Bibi alleged that shortly after he raised concerns about the company’s practices, he was “frozen out” by Unilab’s management, “no longer asked for advice on compliance matters,” and “replaced as General Counsel.” FLPA II, 734 F.3d at 161.

29 Stevens, 529 U.S. at 784 (dismissed because the Supreme Court found that the defendant, an agency of a state government, could not be sued under the statute); Repko I, 490 F. App’x at 503–05 (dismissed because the relator’s disclosure was not considered “voluntary,” as an earlier plea agreement required him to give the government information about the company’s illegal activities).


32 Doe, 862 F. Supp. at 1504.

33 X Corp. II, 816 F. Supp. at 1092. The parties filed two other lawsuits related to this FCA case. The company obtained an injunction prohibiting the lawyer from disclosing confidential information, X Corp. v. Doe (X Corp. I), 805 F. Supp. 1298 (E.D. Va. 1992), and the lawyer filed a wrongful termination counterclaim against the company, Under Seal, 17 F.3d at 1435.

34 X Corp. II, 816 F. Supp. at 1086.
to sue for wrongful termination and provided the company with a copy of his draft complaint. The company preemptively filed a lawsuit against the lawyer, claiming that his planned disclosure of information in his wrongful termination complaint would violate his fiduciary duty and a confidentiality agreement he had signed. Although the company’s lawsuit against the lawyer was based on state (rather than federal) law, the company filed its lawsuit in federal court based on diversity jurisdiction. It asked the court for an injunction requiring the lawyer to return allegedly misappropriated documents and prohibiting him from disclosing any confidential information.

A month later, the lawyer filed his FCA lawsuit in that same federal district court. As required under the FCA, he provided a copy of his complaint and supporting documentation to the Justice Department. The Justice Department became concerned that some of the supporting documentation was subject to the company’s attorney-client privilege and asked the court to hold the FCA lawsuit in abeyance while the company’s lawsuit against the lawyer proceeded.

In the company’s lawsuit against the lawyer, the district court applied the Virginia Code of Professional Responsibility. The Code’s confidentiality rule allowed lawyers to disclose client fraud only if the evidence “clearly establishment[d]” that fraud. The court found that the disputed information was “arguably suggestive of a regulatory violation,” but fell “short of clearly showing fraud.” In response to the company’s lawsuit, the court issued

---

35 Id.
36 Id.
38 See id.
39 Id. The company also alleged that Doe breached his fiduciary duty to the company by revealing confidences to his own attorney. Id. at 1301. The district court rejected the company’s claim that Doe’s disclosures to his own attorney breached his fiduciary obligation because that “would cripple Doe’s ability to defend against X Corp.’s attack on his professional conduct.” Id. at 1301 n.5.
40 See Under Seal, 17 F.3d at 1435.
41 See 31 U.S.C. § 3730(b)(2). The FCA requires relators to serve the government with a copy of the complaint as well as “written disclosure of substantially all material evidence and information the [relator] possesses.” Id.
42 Under Seal, 17 F.3d at 1435.
43 X Corp. I, 805 F. Supp. at 1298 (quoting VA. CODE OF PROF’L RESPONSIBILITY DR 4-101(C)(3) (VA. STATE BAR ASS’N 1983)). (The Virginia Code was in effect when the case was decided.) Virginia’s confidentiality rule permitted the disclosure of a client’s past fraud as long as the evidence “clearly establish[d]” the fraud. Id.
44 Id. The Virginia Code lacked a broad “offensive use” exception analogous to the ABA’s Model Rule 1.6(b)(5). See MODEL RULES OF PROF’L CONDUCT r. 1.6(b)(5); VA. CODE OF PROF’L RESPONSIBILITY DR 4-101.
an injunction prohibiting the lawyer from disclosing this information.\textsuperscript{45} The court eventually dismissed the lawyer-relator’s FCA case based on the earlier injunction.\textsuperscript{46} While rejecting this particular lawyer-relator’s FCA suit, the court nonetheless exhibited solicitude rather than hostility for the concept of a lawyer serving as a relator. It declared: “[T]o the extent that state law permits disclosure of client confidences, such as to prevent a future or ongoing crime or fraud, then the attorney’s use of the \textit{qui tam} mechanism to expose that fraud should be encouraged, not deterred.”\textsuperscript{47}

A second false claims case addressing lawyer confidentiality, \textit{Bury v. Community Hospitals of Central California}, was decided in 2002.\textsuperscript{48} It involved the former General Counsel of a hospital chain, Robert Bury, who sued his former employer under California’s (rather than the federal) False Claims Act four months after the hospital chain terminated him.\textsuperscript{49} The California False Claims statute, like its federal counterpart, requires a “qui tam plaintiff [to] disclose to the Attorney General, in writing, ‘substantially all material evidence and information’ the qui tam plaintiff possesses.”\textsuperscript{50} The court noted the close parallel between this case and \textit{Doe}, and took a similar approach.\textsuperscript{51} The issue in this case was whether Bury’s “duty of confidentiality and loyalty to his former client preclude[d] his \textit{qui tam} complaint,” because he was unable to “legally disclose sufficient information to form the basis of a valid complaint.”\textsuperscript{52} The court indicated that Bury could proceed with his California False Claims lawsuit only if he could “demonstrate that under California law . . . [his] duty of loyalty and confidentiality [did not] prevent[] him from legally disclosing sufficient information to support the complaint.”\textsuperscript{53} California’s confidentiality standard is even stricter than Virginia’s and lacks any exception for client fraud.\textsuperscript{54} Therefore Bury could not pursue his lawsuit; nor, under this logic, could any lawyer subject to California’s rules.

\begin{itemize}
\item \textsuperscript{45} \textit{X Corp. I}, 805 F. Supp. at 1312.
\item \textsuperscript{46} \textit{X Corp. II}, 816 F. Supp. at 1087.
\item \textsuperscript{47} \textit{Doe}, 862 F. Supp. at 1507–08 (footnote omitted).
\item \textsuperscript{48} \textit{Bury}, 2002 Cal. App. Unpub. LEXIS 1035, at *2.
\item \textsuperscript{49} \textit{Id.} Bury’s employment ended on October 30, 1998. \textit{Id.} He filed his qui tam action on February 8, 1999. \textit{Id.}
\item \textsuperscript{50} \textit{Id.} at *6 (quoting \textit{CAL. GOV’T CODE} § 12652(c)(3) (West 2000)).
\item \textsuperscript{51} \textit{Id.} at *7.
\item \textsuperscript{52} \textit{Id.} at *10.
\item \textsuperscript{53} \textit{Id.} at *5.
\item \textsuperscript{54} \textit{Id.} at *8.
\item \textsuperscript{55} See \textit{CAL. RULES OF PROF’L CONDUCT} r. 3-100(B) (\textit{STATE BAR OF CAL. 2015}) (permitting disclosure of confidential client information to prevent criminal acts likely to result in death or substantial bodily harm).
\end{itemize}
A third lawyer-relator case applying lawyer confidentiality standards is United States ex rel. Fair Laboratory Practices Associates v. Quest Diagnostics, Inc., which was decided by the Southern District of New York in 2011 (“FLPA I”) and affirmed by the U.S. Court of Appeals for the Second Circuit in 2013 (“FLPA II”).56 This federal FCA case alleged that from 1996 through 2005, the pricing policy adopted by Unilab, a medical testing company, violated the criminal anti-kickback statute.57 The company’s General Counsel, Mark Bibi, raised concerns about the pricing policy within the company in 1996, and the company adjusted its policy in response.58 But in 1999, new management came in and reinstated the earlier pricing policy.59 After Bibi again raised concerns internally about the policy’s possible illegality, the company removed him as General Counsel.60

In 2005, Bibi and two other former Unilab executives created a corporation, Fair Laboratory Practices Associates (“FLPA”), for the purpose of bringing an FCA lawsuit against their former employer based on its alleged violations of the anti-kickback statute.61 The defendant sought dismissal of the lawsuit, arguing that Bibi violated his confidentiality obligation.62 Bibi argued that his disclosures were permitted under New York’s confidentiality rule,63 which allows a lawyer to disclose “confidential information to the extent that the lawyer reasonably believes necessary . . . to prevent the client from committing a crime.”64 The issue was therefore whether the disclosures that Bibi made in 2005 were permitted under New York’s confidentiality-

57 Id. at *2 (referring to 42 U.S.C. § 1320a-7b(b) as the “Anti-Kickback Statute”).
58 Id. at *10.
59 Id.
60 Id. at *12.
61 Id. at *16.
62 Id. at *17.
63 Id. at *35.
ality rule.\textsuperscript{65} The court found that it was reasonable for Bibi to believe that
the defendant’s violations of the anti-kickback statute were ongoing in 2005.\textsuperscript{66} But Bibi’s disclosures were nonetheless improper because they went
beyond what was necessary to prevent the former client from committing a
crime. In particular, the court found that Bibi’s “disclosure of confidences
from the 1990s to March 2000” was not “necessary to prevent the commis-
sion or continuation of a crime in 2005.”\textsuperscript{67} Because New York permits
disclosures in order to prevent, but not rectify, client crimes, a lawyer-relator
bound by New York rules would be able to reveal information necessary to
stop ongoing crimes or prevent future crimes, but not information about
past crimes. Nor would a lawyer-relator be free to disclose ongoing or fu-
ture frauds that were not criminal under state or federal law.

In addition to cases involving lawyer-relators suing former clients, we
have identified about a dozen cases where a lawyer-relator used information
learned in an earlier representation to sue a non-client third party.\textsuperscript{68} Courts
dismissed most of these cases at an early stage of the litigation without ad-
dressing the lawyers’ ethics obligations, usually because the case was based
on information that had been publicly disclosed and the relator did not qual-
ify as an “original source.”\textsuperscript{69} But in one of these cases, United States ex rel. Hol-
mes v. Northrop Grumman, decided in 2015, the district court explicitly
addressed the lawyer-relator’s confidentiality obligation, finding that he
used the confidential information of a current client without proof that his
client had given informed consent.\textsuperscript{70}

\textsuperscript{65} FLPA I, 2011 U.S. Dist. LEXIS 37014, at *35. The confidentiality exceptions under New
York’s rule are more limited than those in states that follow the Model Rules. See infra notes 78–
123 and accompanying text.

\textsuperscript{66} FLPA I, 2011 U.S. Dist. LEXIS 37014, at *33–34 (“Bibi could have reasonably believed in
2005 that [d]efendants had the intention to commit a crime.”); see also FLPA II, 734 F.3d at 164.

\textsuperscript{67} FLPA I, 2011 U.S. Dist. LEXIS 37014, at *36–37; see id. at *36 (finding relators could
have shown that there was a “continuing crime in 2005” by providing “evidence of Quest’s pric-
ing agreements . . . in effect in 2005”); see also FLPA II, 734 F.3d at 165. The district court also
ruled that Bibi violated Rule 1.9(a), which prohibits subsequent conflicts of interest. FLPA I, 2011
U.S. Dist. LEXIS 37014, at *38. Such a ruling, if it were followed, would be the death knell to
lawyer-relators suing former clients. But the Second Circuit declined to adopt this reasoning, rely-
ong on confidentiality as the basis for dismissal of the suit. FLPA II, 734 F.3d at 165.

\textsuperscript{68} See infra notes 198–241 and accompanying text for a discussion of the loyalty concerns
that arise in this context.

\textsuperscript{69} See, e.g., United States ex rel. Kreindler & Kreindler v. United Techs. Corp., 985 F.2d 1148 (2d Cir. 1993); see also Robert L. Vogel, The Public Disclosure Bar Against Qui Tam Suits, 24 PUBL. CONT. L.J. 477, 517 n.178 (1995) (noting that several of the earliest FCA cases addressing
the “public disclosure bar” were brought by lawyer-relators and that “courts may have been con-
cerned with lawyer/client ‘parasitism,’” i.e., lawyers inappropriately benefiting from information
they learned while representing clients).

\textsuperscript{70} United States ex rel. Holmes v. Northrop Grumman, No. 1:13cv85-HSO-RHW, 2015 U.S.
Donald Holmes represented Munich Re, an insurer, in an arbitration proceeding with Northrup Grumman Corporation (“NGC”), a government contractor. In connection with that representation, Holmes filed a complaint against NGC in federal district court seeking certain government documents for use in the arbitration. He obtained those documents subject to the court’s protective order. He then used the documents in a qui tam FCA lawsuit he filed pro se against NGC alleging that NGC had defrauded the government.

In the FCA lawsuit, the court granted NGC’s motion to disqualify Holmes as relator and to dismiss the complaint, finding that Holmes violated not only the protective order, but also the ethical rules on confidentiality, conflicts of interest, candor to a court, and misrepresentation. With respect to confidentiality, the court found that Holmes breached his duty to keep information related to the representation of Munich Re confidential when he revealed and used for his personal benefit the government documents he had obtained on his client’s behalf. The court also found that although Munich Re had indicated that it did not object to his decision to report NGC’s fraud, Holmes failed to prove that he had obtained Munich Re’s informed consent before revealing its confidential information.

2. Lawyer Confidentiality Exceptions: The ABA Model Rules and State Variations

In the cases discussed above where a lawyer-relator sued a former client, the courts relied on the various confidentiality rules adopted in Virginia, California, and New York. The confidentiality rules found in these three states, however, were somewhat idiosyncratic and stand in contrast to the rules adopted by most states, which more closely track the approach found in the ABA Model Rules of Professional Conduct. This subsection examines how Model Rule 1.6 and some state variations would apply in this context, addressing the issues that were raised in the three cases above as well as other issues that those courts did not address.

71 Id. at *4.
72 Id.
73 Id.
74 Id. at *5.
75 Id. at *33–34.
76 Id. at *26.
77 Id. at *25.
78 See, e.g., id. In Holmes, where the lawyer-relator acted pro se in filing an FCA lawsuit against a non-client third party, the Mississippi district court considered the ethics rules adopted in Mississippi and the District of Columbia as well as the ABA Model Rules in deciding to disqualify the lawyer-relator. See id.
The obligation of confidentiality defined in the ABA Model Rules and adopted in most states is broad in scope, reaching all “information relating to the representation of a client.” But some jurisdictions, including New York, use a narrower formulation based on the earlier ABA Model Code of Professional Responsibility, reaching only information that is subject to the attorney-client privilege (“confidences”), or information that a client has specifically requested to be kept confidential or would be detrimental to the client if revealed (“secrets”). Under either formulation, the lawyer’s obligation continues even after the lawyer-client relationship has ended.

The Model Rule on confidentiality includes several exceptions that are relevant to FCA lawyer-whistleblowers: two distinct but overlapping exceptions addressing client frauds and crime, and an exception for disputes between lawyer and client. One provision, Model Rule 1.6(b)(2), permits disclosure in order “to prevent the client from committing a crime or fraud that is reasonably certain to result in substantial injury to the financial interests or property of another.” A second provision, Model Rule 1.6(b)(3), permits disclosure in order “to prevent, mitigate or rectify” such injury, even if the client’s crime or fraud has already occurred. Because some states (including New York) permit disclosure to prevent client wrongdoing but not to mitigate or rectify it, we must address these two provisions separately.

---

79 Model Rules of Prof’l Conduct r. 1.6(a).
80 See, e.g., D.C. Rules of Prof’l Conduct r. 1.6(a)–(b) (D.C. Bar Ass’n 2015); N.Y. Rules of Prof’l Conduct r. 1.6(a); see also Model Code of Prof’l Responsibility DR 4-101(B)(1) (Am. Bar Ass’n 1983).
81 See Model Rules of Prof’l Conduct r. 1.6. Another confidentiality exception that could come into play applies specifically to organizational clients. Under Model Rule 1.13(b), a lawyer for an organization must engage in internal whistleblowing if the lawyer knows that someone within the organization “is engaged in action . . . that is . . . a violation of law that reasonably might be imputed to the organization, and that is likely to result in substantial injury to the organization.” Under Model Rule 1.13(c), if the organization “insists upon . . . an action . . . that is clearly a violation of law,” and “the lawyer reasonably believes that the violation is reasonably certain to result in substantial injury to the organization,” then the lawyer may reveal confidential information even if Rule 1.6 would not permit it. Under Model Rule 1.13(c). But such disclosure is permitted “only if and to the extent the lawyer reasonably believes necessary to prevent substantial injury to the organization.” Because this confidentiality exception applies only where the disclosure is necessary to prevent injury to the organization, it would not apply to the filing of an FCA lawsuit. While the filing of an FCA lawsuit may be necessary in order to prevent or rectify injury to the financial interests of the United States, it would never be necessary in order to prevent injury to the FCA defendant.
82 Id. r. 1.6(b)(2) (emphasis added). While Model Rule 1.6(b)(2) treats crimes and frauds the same, many states distinguish between crimes and non-criminal frauds, permitting disclosure in order to prevent a client’s crime but not to prevent a non-criminal fraud.
To understand how these two exceptions operate, one must consider three distinct time frames for the crimes or frauds: those that will occur entirely in the future, those that are ongoing, and those that occurred entirely in the past. If a client’s crime or fraud is entirely in the future (and if other criteria are met), then disclosure is permitted under either of these exceptions. But if an FCA violation has not yet occurred, there is no basis for an FCA lawsuit. If a client’s crime or fraud is ongoing (and if other criteria are met), then Rule 1.6(b)(2) permits disclosure to prevent its continuation, and Rule 1.6(b)(3) permits disclosure to mitigate or rectify financial harm that already occurred. If a client’s crime or fraud is entirely in the past, then Rule 1.6(b)(2) does not permit disclosure, but Rule 1.6(b)(3) permits disclosure in order to mitigate or rectify financial harm that has already occurred.

For a lawyer who is considering whether to file an FCA suit, a critical question is whether the applicable confidentiality standard permits disclosure in order to mitigate or rectify the financial harm caused by a client’s past crime or fraud. The New York confidentiality rule permits disclosure in order to stop an ongoing crime, but not to rectify a past one. In FLPA II, the U.S. Court of Appeals for the Second Circuit indicated that the lawyer-relator was allowed to disclose information “necessary to prevent the commission or continuation of a crime in 2005,” when he filed the FCA complaint. But it ruled that he violated New York’s confidentiality rule because he also disclosed “confidences from the 1990s,” disclosures that were not necessary to stop the ongoing crime in 2005. Under the FLPA II court’s analysis, lawyers in New York and similar states may disclose only information that is necessary to stop ongoing criminal FCA violations. In theory, such a limited disclosure could form the basis for an FCA complaint focusing on ongoing violations. But it is not clear whether an FCA complaint limited to ongoing (rather than past) violations could attract a relator’s lawyer, whose compensation is based on the ultimate verdict or settlement, which, in turn, is based on the number and magnitude of the false claims that the defendant filed with the federal government. If there is a company history of filing false claims but the lawyer is ethically prohibited

85 Id. r. 1.6(b)(2)-(3). Model Rule 1.6(b)(3) also permits disclosure in order to prevent future financial harm that is “reasonably certain to result.” See id. r. 1.6(b)(3); see also FLPA II, 734 F.3d at 164–65 (holding that New York’s exception for the prevention of crimes could justify the disclosure of ongoing crimes, but did not justify the disclosure of confidential information about facts that occurred more than five years before FLPA filed the FCA lawsuit).
86 N.Y. RULES OF PROF’L CONDUCT r. 1.6(b)(2) (permitting disclosure “to prevent the client from committing a crime”).
88 Id.
89 See id.
from disclosing all but the most recent, then a relator’s lawyer may be uninterested in filing a lawsuit that would be so limited in scope.

There is significant variation across states regarding confidentiality exceptions related to client crime or fraud. The Model Rules permit disclosure to prevent a client’s crime or fraud only if the client used the lawyer’s services in committing that crime or fraud, but twenty-one states permit such disclosure even if the client did not use the lawyer’s services in the wrongdoing.90 Most states permit disclosure both to prevent future crimes or frauds and to mitigate or rectify past ones.91 If the confidentiality rule in one of these jurisdictions applies, the lawyer-relator will have the most leeway in disclosing a former client’s crime or fraud in the FCA lawsuit. Ten states (including New York) permit disclosure to prevent wrongdoing, but not to rectify or mitigate the harm caused by past wrongdoing. If the confidentiality rule in one of these states applies, the lawyer-relator may be limited in the same way the Second Circuit limited lawyer Mark Bibi in FLPA II: permitting disclosure only to stop ongoing crimes and frauds.92 Fourteen states (including New York) permit disclosure to prevent crimes, but not to prevent non-criminal frauds.93 New Jersey actually requires lawyers to make disclosures that can prevent crimes and frauds as well as illegal acts.94 Six states (Alabama, California, Kentucky, Missouri, Montana, and Rhode Island) lack any confidentiality exceptions for client fraud and monetary crimes—past, ongoing, or future.95 If this kind of restrictive confidentiality

90 MODEL RULES OF PROF’L CONDUCT r. 1.6(b)(2); LATHAM & WATKINS, ATTORNEYS AS SEC WHISTLEBLOWERS: CAN AN ATTORNEY BLOW THE WHISTLE ON A CLIENT AND GET A MONETARY AWARD? 23–28 (2013), http://www.lw.com/thoughtLeadership/SEC-whistleblowers [http://perma.cc/X5ER-Z6T3]. Some of those states also permit disclosure to rectify or mitigate a client’s past crime or fraud, but only if the client used the lawyer’s services in that crime or fraud. As discussed below, disclosure in order to prevent a future fraud is unlikely to form the basis of an FCA lawsuit. See infra note 97 and accompanying text.

91 See infra note 90, at 23–28 (providing a chart of state ethics rules and permitted disclosures); see also ARIZ. RULES OF PROF’L CONDUCT r. 1.6 (STATE BAR OF ARIZ. 2015); HAW. RULES OF PROF’L CONDUCT r. 1.6 (HAW. STATE JUD. 2015); TENN. RULES OF PROF’L CONDUCT r. 1.6 (TENN. BAR ASS’N 2015).

92 See FLPA II, 734 F.3d at 165.

93 See, e.g., N.Y. RULES OF PROF’L CONDUCT r. 1.6(b)(2). The other states that permit disclosure to prevent crimes but not non-criminal fraud are Georgia, Idaho, Kansas, Michigan, Nebraska, New Hampshire, New Mexico, Ohio, Oregon, West Virginia, and Wyoming. Florida and Virginia require disclosure to prevent crime but prohibit disclosure to prevent non-criminal fraud. LATHAM & WATKINS, supra note 90, at 23–28.

94 See N.J. RULES OF PROF’L CONDUCT r. 1.6(b)(1) (N.J. COURTS 2015).

95 ALA. RULES OF PROF’L CONDUCT r. 1.6 (ALA. COURTS 2015); KY. RULES OF PROF’L CONDUCT r. 3.130(1.6) (KY. BAR ASS’N 2015); MO. RULES OF PROF’L CONDUCT r. 4-1.6 (MO. COURTS 2015); MONT. RULES OF PROF’L CONDUCT r. 1.6 (STATE BAR OF MONT. 2011); R.I. RULES OF PROF’L CONDUCT r. 1.6 (R.I. JUDICIARY 2015). All of these states permit disclosure to prevent client crimes that would result in death or serious bodily injury, but not financial crimes.
rule applies, a lawyer-relator may run up against the same barrier that lawyer Robert Bury faced in his unsuccessful state FCA lawsuit. 96

Aside from the temporal dimension, additional complications arise in applying the crime- and fraud-related exceptions. If such an exception applies, a lawyer may disclose only “to the extent the lawyer reasonably believes necessary” to stop the client’s ongoing fraud or crime, 97 or to prevent, mitigate, or rectify injury resulting from the client’s past fraud or crime. 98 Some might contend that filing an FCA complaint is never “necessary”—either to stop a client’s ongoing fraud or to prevent, mitigate, or rectify a past fraud—because a whistleblower could use other means to pursue those goals, such as simply informing the federal government of the client’s alleged fraud. This argument has some force. But in enacting the FCA’s qui tam provisions, Congress arguably has determined that protection of the government’s interest requires not just that whistleblowers be permitted to inform the government of these violations, but also that they be able to pursue an FCA lawsuit on the government’s behalf, even where the government chooses not to participate in the suit. 99 If a whistleblower simply informs the government of a violation, there is no guarantee that the government will devote the resources necessary to investigate the tip—let alone pursue an FCA lawsuit based on it. 100

The qui tam mechanism increases the number of FCA lawsuits by allowing lawyers outside of the Justice Department to bring such suits. 101 It may also increase the quality of information the government receives about FCA

---

97 Model Rules of Prof’l Conduct r. 1.6(b)(2) (emphasis added). Although Model Rule 1.6(b)(2) refers to the prevention of a crime or fraud, it also permits disclosure in order to prevent the continuation of an ongoing crime or fraud. An FCA violation that is entirely in the future (rather than ongoing) cannot form the basis for an FCA lawsuit.
98 Id. r. 1.6(b)(3).
100 See id. Part of Congress’s motivation for the 1986 amendments reviving the qui tam mechanism was concern that the Justice Department was not energetically pursuing FCA lawsuits. See U.S. Gen. Accounting Office, Report by the Comptroller General of the United States: Department of Justice Should Coordinate Criminal and Civil Remedies to Effectively Pursue Fraud in Federal Programs, at i (1979) (“[The Department of] Justice is not making full use of civil remedies to . . . recover losses of program funds due to fraudulent activity.”).
101 See Civil Div., U.S. Dep’t of Justice, supra note 2, at 2. In fiscal year 2014, the Justice Department filed 92 FCA lawsuits and qui tam relators filed 713 suits. Id.
violations in that relators’ lawyers may identify the strongest cases and invest resources in preparing those cases.\textsuperscript{102} Most of the money that the government recovers under the FCA comes from qui tam cases rather than government-initiated FCA lawsuits.\textsuperscript{103} The client intake function at a relators’ law firm—particularly firms that are repeat players with established FCA practices—can serve to identify those cases that are most likely to be financially successful and exclude those that are least likely to result in successful verdicts or settlements.\textsuperscript{104} By the time the government reviews qui tam complaints to decide whether to intervene, relators’ lawyers already have conducted a review, screening out cases that are least likely to succeed.\textsuperscript{105}

If the lawyer is still representing the client, there is an added layer to the “reasonably necessary” analysis. Model Rule 1.4 requires a lawyer to “keep the client reasonably informed about the status of the matter,” and to “explain a matter to the extent reasonably necessary to permit the client to make informed decisions regarding the representation.”\textsuperscript{106} In light of these obligations, before a lawyer engages in external whistleblowing, he or she must communicate with the client about the risks stemming from the client’s violations and how the client can mitigate those risks.\textsuperscript{107} If the lawyer is representing an organizational client and knows that someone “associated with the organization is engaged in action, [or] intends to act” in a way “that is a violation of . . . law that reasonably might be imputed to the organiza-

\textsuperscript{102} See Kwok, \textit{supra} note 17, at 236–37 (noting that when firms spend more time identifying meritorious cases the Department of Justice is more likely to intervene and win).

\textsuperscript{103} See \textit{Civil Div., U.S. Dep’t of Justice, supra note 2, at 2. Of the $44.7 billion recovered under the FCA, more than two-thirds ($30.3 billion) came from qui tam cases. Id.}


\textsuperscript{105} See \textit{The False Claims Act Correction Act: Strengthening the Government’s Most Effective Tool Against Fraud for the 21st Century: Hearing on S. 2041 Before the S. Judiciary Comm., 110th Cong. 412 (2008) (statement of John E. Clark, Of Counsel, Goode, Casseb, Jones, Riklin, Choate & Watson, P.C.) (asserting that relators’ lawyers “choose their cases carefully and always try to choose cases that the Government will . . . intervene in”). Two empirical studies found that experienced relators’ firms cannot be accurately characterized as “filing mill[s]” that “exercise little discretion” in choosing cases and “simply file anything remotely meritorious.” Kwok, \textit{supra} note 19, at 17; see also David Freeman Engstrom, \textit{Harnessing the Private Attorney General: Evidence from Qui Tam Litigation}, 112 COLUM. L. REV. 1244, 1317 (2012). The skill of the relators’ bar in identifying strong cases may be reflected in the fact that most of the money recovered under the FCA comes from qui tam (rather than Justice Department-initiated) cases. On the other hand, the Justice Department intervenes in only 27% of qui tam cases, and most relators’ lawyers voluntarily dismiss cases in which the Justice Department does not intervene. See Kwok, \textit{supra} note 17, at 239–40 (discussing the success rates of intervened and non-intervened qui tam cases).

\textsuperscript{106} \textit{Model Rules of Prof’l Conduct} r. 1.4(a)(3), (b).

\textsuperscript{107} See id. Model Rule 1.4 would require a lawyer to inform the client before filing an FCA lawsuit, but the FCA’s seal provision prohibits such a disclosure. See 31 U.S.C. § 3730(b)(2).
tion, and that is likely to result in substantial injury to it, then the lawyer must engage in internal whistleblowing, which will ordinarily require the lawyer to “refer the matter to higher authority in the organization.” Therefore, before engaging in external whistleblowing to prevent an organizational client from committing—or continuing an ongoing—crime or fraud under Model Rule 1.6(b)(2), the lawyer would first have to engage in internal whistleblowing. If the issue is rectification of an organizational client’s past crime or fraud under 1.6(b)(3), then Model Rule 1.13(b) does not mandate internal whistleblowing, but the Rule 1.4 obligation to keep the client informed would still apply. On the other hand, if the lawyer no longer represents the client, then neither 1.4 nor 1.13(b) would apply.

In addition, these exceptions apply only in situations involving a fraud or crime, and the FCA does not map perfectly onto this requirement for a fraud or crime. The Model Rules define “fraud” as “conduct that is fraudulent under the substantive or procedural law of the applicable jurisdiction and has a purpose to deceive.” An FCA violation, on the other hand, can occur even where a defendant did not specifically intend to defraud the government. A violation can be triggered where the defendant had “reckless disregard of the truth or falsity of the information” it submitted to the government. Therefore, only those FCA cases where the defendant had a purpose to deceive will qualify for the fraud-related exception to confidentiality.

Another confidentiality exception, Model Rule 1.6(b)(5), allows a lawyer to disclose information in order “to establish a claim . . . on behalf of

---

108 See MODEL RULES OF PROF’L CONDUCT r. 1.13(b).
109 See id. r. 1.4, 1.6(b)(3), 1.13(b).
110 See infra notes 197–241 and accompanying text.
111 MODEL RULES OF PROF’L CONDUCT r. 1.0(d).
112 31 U.S.C. § 3729(b)(1). But see United States v. Slocum, 708 F.2d 587, 596 (11th Cir. 1983) (holding that in a criminal prosecution for filing a false claim under 18 U.S.C. § 287, the government must show that the defendant acted with “specific intent to violate the law or with a consciousness that what he was doing was wrong”).
113 31 U.S.C. § 3729(b)(1) (defining the term “knowingly” to include “act[ing] in reckless disregard of the truth or falsity of the information”); see John T. Boese, Of Counsel, Fried, Frank, Harris, Shriver & Jacobson LLP, Paper Presented at the Tenth National Institute on the Civil False Claims Act and Qui Tam Enforcement: Fundamentals of the Civil False Claims Act and Qui Tam Enforcement, at A-11 (June 4, 2014) (discussing the “reckless disregard” standard) (on file with authors).
114 These confidentiality exceptions apply only in situations involving “substantial injury” to someone’s “financial interests.” See MODEL RULES OF PROF’L CONDUCT r. 1.6(b)(2)–(3). This requirement is likely to be met in qui tam lawsuits under the FCA because relators must be represented by counsel who are generally paid on a contingent basis. Relators’ counsel are unlikely to take on such representation unless the case has the potential for significant financial returns, so FCA cases will generally satisfy the “substantial injury” requirement of the crime- and fraud-related exceptions.
the lawyer in a controversy between the lawyer and the client." The law-
ner-relator’s FCA lawsuit against a client is arguably “a controversy be-
 tween the lawyer and the client,” both formally (because the relator is a par-
dy to the lawsuit) and in substance (because the U.S. Supreme Court has
ruled that the FCA “effect[s] a partial assignment of the Government’s dam-
egages claim” to the relator). On the other hand, some may view the situa-
tion—at least up until the FCA lawsuit is filed—as an inchoate controversy
between the government and the company, rather than an actual controversy
between the lawyer and the company. It is the filing of the FCA lawsuit that
creates the partial assignment. Up until that point, the prospective relator is
merely a potential witness to an inchoate dispute between the government
and the prospective FCA defendant (rather than a party to a dispute between
the government and the FCA defendant). Judges may look skeptically upon
a lawyer who tries to use the 1.6(b)(5) confidentiality exception to justify
the disclosures necessary for filing an FCA lawsuit. Most states have adopt-
ed the current formulation of Model Rule 1.6(b)(5), but three jurisdic-
tions—the District of Columbia, Michigan, and New York—limit a lawyer’s
ability to use confidential information offensively against a client to situa-
tions where the lawyer is attempting to establish or collect unpaid legal fees
(rather than pursuing other claims against the client). California lacks any
express exception for disputes between lawyers and their clients.

The discussion above focuses on the confidentiality obligations of a
lawyer who wishes to sue a current or former client. In a different context—
where a lawyer wishes to sue a non-client third party—the confidentiality
analysis differs. In that situation, the lawyer’s ability to disclose information
depends on whether the applicable rule employs the narrower standard, reach-

\[\text{References}\]

115 Model Rules of Prof’l Conduct r. 1.6(b)(5). Lawyer-whistleblowers who have expe-
vanced retaliation often invoke this exception when they sue former clients and seek whistleblower
protection. See, e.g., Burkhart v. Semitool, Inc., 5 P.3d 1031 (Mont. 2000); ABA Standing
Comm. on Ethics and Prof’l Responsibility, Formal Op. 01-424 (2001) (opining that a former in-
house lawyer may pursue a wrongful discharge claim against a former employer and client as long
as client information is properly protected).

116 Stevens, 529 U.S. at 773.

117 See John M. Barkett, Partner, Shook, Hardy & Bacon LLP, Paper Presented at the American
Bar Association, Section of Litigation Annual CLE Conference: Lawyer-Client Fallout: Using
Privileged Information to Establish a Claim or Defense Against a Client/Employer, at 5 (Apr.
2013), http://www.americanbar.org/content/dam/aba/administrative/litigation/materials/sac2013/sac_
most states have adopted the text of Model Rule 1.6(b)(5)); see also Model Rules of Prof’l
Conduct r. 1.6(b)(5). The earlier ABA Model Code of Professional Responsibility included a
similar exception, but it was narrower in scope, permitting disclosure only to establish or collect a

118 D.C. Rules of Prof’l Conduct r. 1.6(e)(5); Mich. Rules of Prof’l Conduct r.
1.6(e)(5)(State Bar of Mich. 2015); N.Y. Rules of Prof’l Conduct r. 1.6(b)(5).
ing only information that is a “confidence” (i.e., the lawyer learned it through a privileged communication from the client) or a “secret” (i.e., the client specifically requested that it be kept confidential or disclosure of it would be detrimental to the client),119 or the broader standard, reaching all “information relating to the representation of a client.”120 Under the narrow formulation, information about a third party’s FCA violation would not even be covered by the confidentiality duty unless the client conveyed the information to the lawyer in a confidential communication, the client specifically requested that it be kept confidential, or disclosure would be detrimental to the client.121 Under the broader formulation, the information is likely to be covered by the confidentiality duty, and the confidentiality exceptions discussed above for client crimes and frauds and lawyer-client disputes would not apply because it is a third party—rather than a client—that is involved in the crime, fraud, or dispute.122 The lawyer who wishes to disclose such information in an FCA lawsuit against a non-client third party would need to obtain the client’s informed consent before making the disclosure.123

C. Lawyers’ Loyalty Obligations and the FCA

Until recently, any discussion of ethical restrictions on a lawyer attempting to take advantage of whistleblower bounties under the False Claims Act was confined to the lawyer’s duty of confidentiality, with no discussion of the lawyer’s duty of loyalty. In Holmes, however, a federal district court disqualified a lawyer-relator who took positions in an FCA lawsuit against a non-client third party that were in conflict with positions he was taking on behalf of a current client in separate litigation, despite the client’s manifested lack of objection to the lawyer’s conduct.124 And in FLPA I, a federal district court held that a partnership that included the company’s former General Counsel was barred from serving as a qui tam relator, regardless of whether the former lawyer had impermissibly disclosed confidential client information, because his role in the lawsuit impli-

---

119 See, e.g., N.Y. RULES OF PROF’L CONDUCT r. 1.6(a).
120 See, e.g., Mo. RULES OF PROF’L CONDUCT r. 1.6(a). This would reach information the lawyer learned in the course of representing a client. See RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 59 cmt. c (AM. LAW INST. 2000).
121 See, e.g., N.Y. RULES OF PROF’L CONDUCT r. 1.6(a).
122 See, e.g., Mo. RULES OF PROF’L CONDUCT r. 1.6(a). On the other hand, one state, Massachusetts, permits disclosure in order to prevent a criminal or fraudulent act by a non-client. MASS. RULES OF PROF’L CONDUCT r. 1.6(b)(1) (MASS. SUP. JUD. CT. 2015).
123 See, e.g., Holmes, 2015 U.S. Dist. LEXIS 71804, at *25–26. Unlike the confidentiality exceptions relating to client fraud, the exception for client consent does not vary significantly from state to state.
124 Id. at *31–32.
icated him in an impermissible conflict of interest with his former client in violation of the New York Lawyer’s Code of Professional Responsibility.125

We begin by discussing a lawyer’s loyalty obligations to current clients and then consider a lawyer’s loyalty obligations to former clients. These loyalty obligations include disciplinary rules concerning conflicts of interest and keeping the client informed of developments material to the representation, as well as state fiduciary law that prohibits lawyers from financially profiting as a result of using or disclosing confidential client information, even when the lawyer is ethically permitted to use or disclose the information for some other purpose.

1. Loyalty Obligations to a Current Client

The most glaring risk of a loyalty violation presumably would involve a lawyer-relator filing an FCA lawsuit against a current client. In all of the reported FCA cases involving a lawyer blowing the whistle on a client company, however, the lawyer had left the company before filing a qui tam complaint.126 It may be unlikely that a lawyer would file a qui tam lawsuit against a current client, particularly an employer: although the complaint is initially filed under seal, the company will learn that the lawyer has filed the complaint as soon as the seal is lifted, making it difficult for the lawyer to continue representing the company.127 Nevertheless, it is possible that a lawyer could file a qui tam complaint against a current client. Indeed in Doe,128 the lawyer claimed that he was contemplating filing a qui tam action against his employer and that he had clandestinely begun to copy confidential documents before he was terminated, allegedly in response to his activity.129 Whether or not he would actually have filed the complaint while still employed by the company is unclear; however, it is certainly foreseeable that a lawyer who is planning to leave, or believes that he or she might be terminated, will begin preparing for a subsequent qui tam filing, includ-

126 See id. at *13; see also Stevens, 529 U.S. at 770; Repko I, 490 F. App’x at 503–04; Doe, 862 F. Supp. at 1504; cf. Bury, 2002 Cal. App. Unpub. LEXIS 1035, at *3 (lawsuit brought under the state version of the FCA).
127 Both state and federal anti-retaliation laws provide lawyers with a basis to argue that they are legally protected against retaliation for engaging in protected whistleblower activity. See infra notes 302–351 and accompanying text. The situation is quite different under Dodd-Frank, where the whistleblower is permitted to disclose information to the SEC anonymously, and the SEC may not reveal the whistleblower’s identity even after a whistleblower award is made. See infra notes 242–301 and accompanying text.
129 Id. at 1096. The court rejected the lawyer’s retaliation claim and did not discuss any possible conflict of interest. Id.
ing locating and copying documents necessary to support the complaint. This preparation activity itself raises a conflicts issue concerning a current client.\textsuperscript{130}

Not all FCA cases involving lawyer-relators concern lawsuits against a client—whether current or former. Some involve lawyers who obtained information about a third party while representing a client, typically in a litigation matter, and then used that information as the basis for filing a qui tam lawsuit against the third party.\textsuperscript{131} In most of these cases the courts did not mention any potential loyalty issue involving the client. In one case, the court briefly identified a potential conflict of interest, but even then, the reference was merely to indicate that dismissal of the lawyer-relator’s case on standing grounds worked “no ‘technical’ or unfair result.”\textsuperscript{132} In Holmes, the court readily identified a conflict of interest based on inconsistent positions the lawyer was taking in the FCA lawsuit against a third party and in related litigation involving the current client.\textsuperscript{133} In that case, the client told the Justice Department that it had no objection to the lawyer’s conduct, but the court found no evidence of the client’s informed consent.\textsuperscript{134} In other cases, the conflict of interest might be more subtle, and even when there is no conflict of interest, clients in future cases might well protest that the lawyer stole an opportunity that should have been presented to the client, thereby acting in violation of applicable ethical standards.

\textit{a. Serving as a Qui Tam Relator Against a Current Client}

Although it is unusual, lawyers sometimes sue a current client; for example, when a lawyer sues a client for unpaid legal fees\textsuperscript{135} or when an in-house lawyer sues his or her employer for violating an anti-discrimination law.\textsuperscript{136} Thus, although unlikely, it is not inconceivable that a lawyer will file a qui tam complaint while still employed by the defendant company. Indeed, anti-retaliation laws provide the lawyer with a basis to argue that, as

\textsuperscript{130}See infra notes 135–169 and accompanying text.


\textsuperscript{132} Fed. Recovery Servs., 72 F.3d at 453.

\textsuperscript{133} Holmes, 2015 U.S. Dist. LEXIS 71804, at *14–16.

\textsuperscript{134} Id.

\textsuperscript{135} See, e.g., In re Simon, 20 A.3d 421 (N.J. 2011); In re Disciplinary Action Against Szymialis, 557 N.W.2d 543 (Minn. 1997).

\textsuperscript{136} See, e.g., Jones v. Flagship Int’l, 793 F.2d 714 (5th Cir. 1986); St. John v. Emp’t Dev. Dep’t, 642 F.2d 273 (9th Cir. 1980).
disloyal as such an act may appear to the company, the lawyer is legally protected against retaliation for having filed the complaint.\textsuperscript{137} Does a lawyer who serves as a qui tam relator against a current client have an impermissible conflict of interest under state rules of professional conduct? For the purpose of addressing this question, we will assume that the lawyer has permission under state confidentiality rules to disclose the client’s information in serving as a qui tam relator.

Ordinarily, a lawyer-relator will have learned information concerning the client’s allegedly illegal conduct as a result of representing the client with respect to the subject matter of the qui tam lawsuit.\textsuperscript{138} If that representation is ongoing at the time the lawsuit is filed, then the lawyer almost certainly has a conflict of interest under Model Rule 1.7, in which a concurrent conflict exists whenever “there is a significant risk that the representation of one or more clients will be materially limited by . . . a personal interest of the lawyer.”\textsuperscript{139} Because the prospect of receiving a large sum of money “might tend to cloud a lawyer’s professional judgment,”\textsuperscript{140} such a risk is clearly present when the lawyer is simultaneously advising the company as to whether it is violating the law, whether the legal violation poses a threat to the company, or whether suspected wrongdoing should be reported to a higher level in the company, including the board of directors.\textsuperscript{141} Such a risk is also present when the lawyer is conducting or monitoring a compliance effort. All these activities require both an objective analysis of the company’s legal obligations and an objective weighing of alternative courses of action available to the client, and it is difficult to expect objectivity from a lawyer who has filed a lawsuit in which the lawyer’s recovery depends on a finding that the client engaged in illegal conduct.\textsuperscript{142}

Under Model Rule 1.7(b), lawyers may accept or continue a representation burdened with a conflict of interest if the lawyer “reasonably believes that the lawyer will be able to provide competent and diligent representation

\begin{footnotes}
\item[137] See \textit{X Corp. II}, 816 F. Supp. at 1095–96 (citing 31 U.S.C. § 3730(h)).
\item[138] Under the ABA Model Rules, the lawyer would not be permitted to use or disclose the information to prevent, mitigate, or rectify a client’s crime or fraud unless the client has used or is using the lawyer’s services in furtherance of the crime or fraud. MODEL RULES OF PROF’L CONDUCT r. 1.6(b)(3); see \textit{supra} notes 26–77 and accompanying text.
\item[139] MODEL RULES OF PROF’L CONDUCT r. 1.7(a)(2).
\item[140] \textit{N.Y. Cnty. Lawyers’ Ass’n}, \textit{supra} note 8, at 10–11 (addressing concurrent conflicts under Dodd-Frank).
\item[141] Cf. Temkin & Moskovits, \textit{supra} note 8, at 21 (explaining that a lawyer who prematurely blows the whistle may harm the client because financial incentives place the lawyer’s personal interests in conflict with the client’s interests).
\item[142] Cf. \textit{id}. The size of the attorney’s recovery may depend on how long the illegal conduct continued, thereby giving the lawyer a financial incentive not to vigorously press the client to stop any illegal conduct.
\end{footnotes}
to [the] client” and the client gives informed consent. 143 If the potential award is large, it is probably unreasonable for the lawyer to believe that he or she would be able to provide “competent and diligent representation.” 144 But in any case, a lawyer-relator could not obtain a client’s informed consent because the FCA prohibits the relator from disclosing the qui tam lawsuit until the case has been filed, the Justice Department has investigated, and the seal has been lifted. 145 Of course, even if the lawyer could inform the client of the qui tam lawsuit, it is difficult to imagine a client permitting its lawyer to serve as a relator in an FCA lawsuit against the client.

As a result, state conflict of interest rules—which, unlike confidentiality rules, do not vary significantly among jurisdictions—apparently prohibit a currently employed lawyer-relator from continuing to work on a matter that is the subject of the qui tam complaint. But perhaps the lawyer has ceased working on that matter or has requested reassignment after filing the qui tam lawsuit. 146 Or perhaps the lawyer never represented the client on that matter, but learned of the illegality as a result of working on some other matter. 147 Putting aside any possible duty to the employer as a former client (with respect to those matters on which the lawyer previously worked but is no longer working), 148 does a lawyer owe a duty to a current client not to sue it, even in an unrelated matter?

In addition to “material limitation” conflicts, Model Rule 1.7 also provides that a concurrent conflict exists whenever “the representation of one client will be directly adverse to another client” even when the matters are entirely unrelated. 149 But this rule expressly applies only when the lawyer will be directly adverse to a client on behalf of another client, not on the lawyer’s own behalf. Arguably, lawyers should not be permitted to do directly, as parties, what they cannot do indirectly, as counsel for a party. 150 If so, then lawyers would not be permitted to take directly adverse action against a current client on behalf of themselves when they could not do so on behalf of another client. This position is supported by a minority of state

---

143 MODEL RULES OF PROF’L CONDUCT r. 1.7(b)(1).
145 See supra notes 11–25 and accompanying text (discussing FCA claim procedures).
146 Cf. Green & Thomas, supra note 144.
147 The ABA Model Rules permit disclosure only if the client had used or was using the lawyer’s services in furtherance of the crime or fraud, but some state rules permit disclosure without any such restriction. See supra notes 26–77 and accompanying text.
148 See infra notes 197–241 and accompanying text (discussing former-client conflicts).
149 MODEL RULES OF PROF’L CONDUCT r. 1.7(a)(1) (emphasis added).
150 See id.; see also FLPA I, 2011 U.S. Dist. LEXIS 37014, at *30–32.
rules that incorporate the former ABA Model Code’s provision that a lawyer may not “intentionally [p]rejudice or damage his client during the course of the professional relationship [except when expressly permitted to do so].” But that broad proposition was not incorporated into the Model Rules that limit the lawyer’s duty of commitment and zeal to those matters for which the lawyer is actively representing the client.

In addition to Rule 1.7, Model Rule 1.8 addresses certain specific conflicts of interest with current clients. Rule 1.8(a) provides that “[a] lawyer shall not enter into a business transaction with a client or knowingly acquire an ownership, possessory, security or other pecuniary interest adverse to a client” unless certain criteria are met, including obtaining the client’s informed consent. The text of this rule would seem to prohibit a lawyer from serving as an FCA relator against a client. By filing an FCA lawsuit, a relator obtains a partial assignment of the government’s FCA claim against a defendant, thus arguably acquiring a “pecuniary interest adverse to a client.” And the FCA’s seal requirement prohibits a relator from obtaining the informed consent of the defendant.

Model Rule 1.8(a) is typically applied, however, to face-to-face transactions between lawyers and their clients, even when the lawyer is not representing the client in that transaction. Indeed, the first comment to Rule 1.8(a) is captioned “Business Transactions Between Client and Lawyer” and indicates that the purpose of the rule is to protect the client against the possibility of overreaching as a result of the “lawyer’s legal skill and training, together with the relationship of trust and confidence between lawyer and client.”

There is no transaction between lawyer and client when a lawyer

---

151 See, e.g., D.C. RULES OF PROF’L CONDUCT r. 1.3(a)(2).
152 See MODEL RULES OF PROF’L CONDUCT r. 1.3 (“A lawyer shall act with reasonable diligence and promptness in representing a client.”); cf. 2 GEOFFREY C. HAZARD & W. WILLIAM HODES, THE LAW OF LAWYERING: A HANDBOOK ON THE MODEL RULES OF PROFESSIONAL CONDUCT § 55.05 illus. 55-1 (4th ed. 2015) (asserting that a lawyer serving on the board of a legal services organization that voted in favor of filing a class action against a bank represented by the lawyer would not violate Model Rule 1.7, but Model Rule 6.3 requires that lawyer to abstain from discussing or voting on the issue).
153 MODEL RULES OF PROF’L CONDUCT r. 1.8(a) (emphasis added).
154 See Stevens, 529 U.S. at 765–66.
155 MODEL RULES OF PROF’L CONDUCT r. 1.8(a). Professor Anthony Sebok suggests that the use of the word “acquires” in the text of the rule would not clearly apply to an FCA relator: although the relator has a pecuniary interest in the success of the qui tam lawsuit, the relator has not “acquired” that interest in the ordinary usage of that word. Anthony Sebok, Professor, Benjamin N. Cardozo Sch. of Law, Comments at the Legal Ethics Scholars’ Roundtable (Oct. 31, 2014).
156 See 31 U.S.C. § 3730(b)(2); supra notes 143–145 and accompanying text (discussing Model Rule 1.7’s informed consent exception).
157 See MODEL RULES OF PROF’L CONDUCT r. 1.8(a) cmt. 1. None of the other comments applicable to 1.8(a) address a situation in which a lawyer acquires a pecuniary interest adverse to a
seeks a whistleblower bounty, and some may argue that the rule should not apply in this situation. If it does not, then lawyer conduct rules might not prohibit lawyers from seeking a whistleblower reward by filing a qui tam lawsuit against a current client, so long as the lawyer is not currently representing the client with respect to the subject matter of the lawsuit. 158

So far we have been positing a lawyer filing a qui tam lawsuit against a current client. What may be more likely, however, is that a lawyer who is planning to leave the company, either voluntarily or involuntarily, will restrict his or her activity to preparing to file a qui tam lawsuit; for example, by clandestinely gathering evidence of the company’s illegal conduct. 159

May a lawyer do so and avoid violating state professional conduct rules by waiting until leaving the company to actually file the complaint? 160

If the lawyer is representing the client on the matter, then preparing to file a qui tam lawsuit almost certainly involves a conflict of interest for the lawyer, for the same reasons identified when the lawyer has actually filed the lawsuit. But when does the conflict arise? Does it arise when the lawyer begins gathering and copying documents? Would it make a difference if the lawyer is contemplating the filing of a lawsuit but has not yet made the decision to do so, or does the mere contemplation of the filing of such a lawsuit create a material limitation conflict? It is conceivable that a lawyer who has been urging greater compliance efforts, perhaps even advising the client of the possibility of a qui tam lawsuit by other employees, will entertain the thought of filing such a lawsuit him or herself, particularly if the lawyer is concerned that a client is hostile to the lawyer’s advice and might fire the lawyer for continuing to press the matter. Must the lawyer then immediately withdraw from the representation or inform the client that he or she has considered the possibility of filing a qui tam lawsuit? We doubt that the mere possibility of the lawyer serving as a qui tam relator will constitute a “significant risk of a material limitation,” 160 although the closer the lawyer comes to the decision to file (including a decision to file if and when the

---

158 If the lawyer previously represented the client with respect to the subject matter of the FCA lawsuit, then the former-client conflict rule may apply. See infra notes 197–241 and accompanying text.

159 This is precisely what the lawyer did in Doe. See 862 F. Supp. at 1504 (noting that “Doe took with him approximately 4300 copies of X Corp. documents and files” when he left X Corp.).

160 See MODEL RULES OF PROF’L CONDUCT r. 1.7(a)(1); supra notes 135–151 and accompanying text.
lawyer is terminated or voluntarily leaves the company), and the more the lawyer does to prepare for such a filing, the more likely it is that the lawyer will be violating the current conflicts rule.

In addition to conflict of interest rules, Model Rule 1.4 requires lawyers to “keep a client reasonably informed about the status of a matter.” As a result, lawyers who currently represent a company in a matter have a duty to inform the company if they are aware that a qui tam action has been or is going to be filed. But the FCA requires that the complaint be filed under seal, which would prohibit a lawyer-relator from either providing the company with a copy of the complaint or informing it of the substance of the complaint. Thus lawyer-relators who comply with the FCA’s requirement of maintaining the confidentiality of the qui tam complaint will violate their ethical duties to the company under Model Rule 1.4. If, however, the lawyer never represented the company on a related matter, or has ceased representing the company on that matter, then there may be no obligation to inform the company, and the lawyer could comply with the qui tam seal requirements without violating this rule.

Finally, even when no disciplinary rule is violated, lawyers have common-law fiduciary duties that prohibit disloyalty to both current and former clients. Do lawyers violate their common-law fiduciary duties when they seek to profit from the use or disclosure of client information, even when the disclosure itself is permitted? Section 60(2) of the Restatement (Third) of the Law Governing Lawyers prohibits such self-dealing: except with the client’s consent, “a lawyer who uses confidential information of a client for the lawyer’s pecuniary gain other than in the practice of law must account to the client for any profits made.” Comment j of section 60(2) makes clear that this fiduciary duty is broader than the prohibitions provided in lawyer disciplinary codes and is derived from the law of agency, under which an agent “has a duty to account for any profits made by the use of such [client] information,” even when the use “does not harm the princi-
Indeed, under the law of agency, the prohibition against self-dealing applies even when the information used "does not relate to the transaction in which he is then employed," and even when the lawyer’s use or disclosure of the information is not itself improper. The purpose of the prohibition is to prevent fiduciary agents from taking personal advantage of client information and to protect against risk to a principal’s interest that may arise when an agent pursues material benefits from third parties in connection with actions taken on behalf of the principal. Although the rule is not reflected in states’ current disciplinary rules, it appears that in seeking a whistleblower bounty under the FCA, lawyers violate their fiduciary duty not to personally profit at their clients’ expense. Indeed, unless the FCA preempts this fiduciary duty, it apparently precludes lawyers from ever serving as qui tam relators against their current or former clients, regardless of the circumstances and regardless of whether they are permitted to disclose adverse information under an exception to the confidentiality rule.

---

164 Id. § 60 cmt. j (citing RESTATEMENT (SECOND) OF AGENCY § 388 cmt. c (AM. LAW INST. 1958)). The Restatement (Third) of the Law Governing Lawyers provides:

Subsection (2) prohibits a lawyer from using or disclosing confidential client information for the lawyer’s personal enrichment, regardless of lack of risk of prejudice to the affected client. The duty is removed by client consent . . . . The sole remedy of the client for breach of the duty is restitutionary relief in the form of disgorgement of profit (see Restatement Second, Agency § 388, Comment c . . . . The strict confidentiality duty of the Subsection is warranted for prophylactic purposes. A lawyer who acquires confidential client information as the result of a representation should not be tempted by expectation of profit to risk a possibly incorrect assessment of future harm to a client. There is no important societal interest in permitting lawyers to make unconsented use or revelation of confidential client information for self-enrichment in personal transactions.

Id. (emphasis added).

165 RESTATEMENT (SECOND) OF AGENCY § 395; see also RESTATEMENT (THIRD) OF AGENCY § 8.02 cmt. b (AM. LAW INST. 2006).

166 RESTATEMENT (SECOND) OF AGENCY § 395 cmt. e ("Even though the agent properly acquires and uses confidential information concerning his principal’s activities in the course of employment, he has a duty to account to the principal for any profits thereby made.").

167 See RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 60(2) cmt. j.

168 RESTATEMENT (THIRD) OF AGENCY § 8.02 cmt. b (discussing the “ordinary expectation that a person who acts as an agent does so to further the interests of the principal and that it is the principal who should benefit from turns of good fortune that may occur in connection with transactions that the agent undertakes on the principal’s behalf”).

169 The failure of the Model Rules and state disciplinary rules to codify all of a lawyer’s fiduciary duties is not inadvertent. There are many instances in which rule drafters make a deliberate choice to limit the types of conduct for which lawyers are subject to discipline, understanding that there are other avenues for enforcement of broader legal duties.
b. Serving as a Qui Tam Relator Against a Third Party

Not all FCA cases involving lawyer-relators are lawsuits against either current or former clients. Some arise when lawyers obtain information about a third party while representing a client, typically in a litigation matter, and then use that information as the basis for a qui tam lawsuit against the third party. In only two of these cases did the court even mention a potential conflict of interest involving the client,\(^{170}\) but we believe that a conflict of interest will often be present in these cases, either because the lawyer’s personal interest in a potential qui tam award may affect the lawyer’s representation of the current client or because the lawyer has taken advantage of an economic opportunity that should have been presented to the client.

In *Holmes*, the lawyer-relator obtained information about NGC while representing Munich Re, an insurer, in an arbitration proceeding with NGC, its insured.\(^{171}\) At the same time that Holmes was representing Munich Re in the arbitration, he filed an FCA complaint against NGC in which he alleged that NGC’s fraudulent conduct duped the government into paying funds to NGC as compensation for damages during Hurricane Katrina.\(^{172}\) The court found that this position conflicted with the position he took as counsel for Munich Re in the arbitration, in which Munich Re argued that it did not owe compensation to NGC for Katrina-related losses because the government had previously paid NGC for those losses.\(^{173}\) The court found a material limitation conflict of interest under ABA Model Rule 1.7, as well as the analogous rules in the District of Columbia and Mississippi.\(^{174}\) The court also rejected Holmes’s claim that he had obtained consent from Munich Re, finding that “the timing of Munich Re’s ‘consent’ [was] questionable and there [was] no evidence the consent was ‘informed.’”\(^{175}\) Although Munich Re was not complaining about Holmes’s conduct, NGC raised the conflict of interest as one of several bases for disqualifying Holmes from serving as relator and dismissing the complaint, and the court agreed.\(^{176}\)

Although the *Holmes* court readily found a conflict of interest based on the conflicting positions Holmes took in two separate proceedings, whether a conflict exists between a current client and a third-party FCA defendant will not always be obvious. To illustrate a more typical case, consider a se-

---

172 Id. at *4.
173 Id. at *18–19.
174 Id. at * 21.
175 Id.
176 Id. at *17–22.
ries of cases involving a single law firm, Stinson, Lyons, Gerlin & Bustamante (“Stinson”) that represented a client who filed a personal injury lawsuit following an automobile accident. In the course of that representation, the law firm came to believe that the defendant’s insurance carrier, Provident, was filing claims in violation of federal law. When the insurance carrier filed a declaratory judgment action against Stinson’s client, seeking to establish the legality of its claims practices, the law firm obtained information in discovery revealing that other insurance companies were involved in similar claims processing. Ultimately, Stinson filed a series of qui tam lawsuits in six different federal district courts naming five different insurance companies as defendants.

In none of the reported decisions involving the Stinson firm did anyone raise the prospect that the Stinson lawyers may have had a conflict of interest concerning their decision to file as a qui tam relator in lawsuits against insurance companies that were engaged in illegal claims processing. Under Model Rule 1.7, however, there would have been a conflict of interest if the lawyers’ personal interest in the qui tam lawsuit—against either Provident or the other insurers—presented a significant risk of materially limiting Stinson’s representation of its client. If Stinson filed a qui tam lawsuit against Provident prior to the conclusion of its client’s personal injury lawsuit, then there may well have been such a significant risk: Prudential’s consent was necessary to settle the lawsuit (including settling the case while any trial judgment was under appeal), and Provident could easily have been so offended by the FCA lawsuit that it would be disinclined to settle the client’s personal injury lawsuit against its insured on terms that were favorable to the client (and thereby to the law firm). There is perhaps less risk of material limitation with respect to an FCA lawsuit.


178 See Provident I, 721 F. Supp. at 1248.

179 See Prudential, 944 F.2d at 1151.

180 This was probably because the law firm’s client was not a party to any of the lawsuits. It is unclear what a court would or should have done if the conflict had been raised by one of the defendant companies. The primary issue raised in the reported opinions concerning the Stinson lawsuits was whether Stinson was an original source of the information upon which the qui tam complaints were based. See, e.g., Chandler, supra note 177, at 550–54.
against the other insurers; however, even those lawsuits posed at least an indirect risk to Provident that could adversely affect the law firm’s representation of the personal injury plaintiff.\textsuperscript{181} Although it is unclear whether the other insurer defendants in the FCA lawsuit could raise such a conflict,\textsuperscript{182} it could certainly be used as the basis for either a subsequent disciplinary action against the lawyer or a subsequent breach of fiduciary duty lawsuit brought by the personal injury client.

Most conflicts arising under Model Rule 1.7 are consentable, and there is no obvious reason why a conflict arising under that rule in cases similar to those involving Holmes and Stinson could not be cured by obtaining the client’s informed consent. Does informed consent in such a case necessarily require advising the client that the client could serve as the qui tam relator? Arguably not, because whether the client might take advantage of this business opportunity may have no bearing on any risk to the client of the lawyer continuing the representation with a personal interest conflict. But for the client’s “consent” to be “informed,” a lawyer must “communicate[] adequate information” not just “about the material risks of . . . the proposed course of conduct,” but also “reasonably available alternatives.”\textsuperscript{183} Perhaps one of those alternatives is that the client could serve as the qui tam relator.\textsuperscript{184}

None of the courts deciding these qui tam lawsuits brought by lawyers against third parties has directly addressed a lawyer’s obligation to advise the client of the opportunity to serve as a qui tam relator. In 1996, the U.S. Court of Appeals for the Fifth Circuit came close to doing so in Federal

\textsuperscript{181} Because Provident used the same allegedly fraudulent claims processing as the other insurers, it could anticipate subsequently being sued by another qui tam relator or by the government itself, even if the law firm had not filed a qui tam lawsuit against Provident. Provident might have been angry with Stinson for publicly airing the fraud allegations and therefore might be disinclined to enter into a favorable settlement with the personal injury client.

\textsuperscript{182} The Holmes court was not troubled by the fact that it was NGC and not Munich Re that was complaining about the conflict of interest between Holmes and Munich Re, but other courts have refused to disqualify a lawyer when the conflict does not adversely affect the complaining party. See, e.g., Universal City Studios, Inc. v. Reimerdes, 98 F. Supp. 2d 449, 455 (S.D.N.Y. 2000) (denying a motion to disqualify when the plaintiff made no effort to show its interests would be adversely affected by the law firm’s conflict); Gilbert v. Knoxville Int'l Energy Exposition, 547 F. Supp. 53, 54 (E.D. Tenn. 1982) (denying a motion to disqualify when the complaining party failed to show any interest that could be adversely affected by the representation).

\textsuperscript{183} MODEL RULES OF PROF’L CONDUCT r. 1.0(c).

\textsuperscript{184} Model Rule 1.7 may not apply because any risk that the lawyer’s personal interest will affect the representation of the client is not significant. MODEL RULES OF PROF’L CONDUCT r. 1.7. Rule 1.4 requires the lawyer to keep the client reasonably informed concerning the representation, but it is unclear whether the scope of the “representation” as to which the client must be informed includes the opportunity to file a qui tam lawsuit against a third party. See MODEL RULES OF PROF’L CONDUCT r. 1.4.
Recovery Services, Inc. v. United States. There, the law firm’s client, Priority E.M.S., Inc. (“Priority”), initially sued its competitor, Crescent City E.M.S., Inc. (“Crescent City”), for engaging in unfair trade practices by filing fraudulent claims with the federal government for reimbursement for ambulance services. The law firm and the President of Priority, Michael Boatright, incorporated Federal Recovery Services (“FRS”) for the express purpose of serving as a qui tam relator in a lawsuit against Crescent City under the FCA. The law firm controlled a majority of the shares of FRS. The government intervened in the FCA case, settled with the defendant, and agreed to pay Boatright ten percent of the settlement as relator’s share. But the government refused to pay the law firm any portion of the settlement, and the court refused to award it the attorney fees ordinarily awarded to lawyers of successful relators. The Fifth Circuit affirmed the dismissal of FRS as a relator because the information had been publicly disclosed in earlier litigation; thus, FRS was not the “original source” for that information. Judge Patrick Higginbotham noted that “the attorneys bypassed a suit by Boatright, their client, in favor of an entity they controlled,” and characterized their actions as an “overreach,” stating that the FCA “did not dispense with the tradition that a lawyer must represent his client’s interest, not his own.” Judge Higginbotham cited no authority for these statements, however, and did not explain how the law firm violated its obligation of loyalty to its clients.

In addition to Model Rule 1.7, Rule 1.8(b) provides that “[a] lawyer shall not use information relating to the representation of a client to the disadvantage of the client unless the client gives informed consent, except as permitted or required by these Rules.” In circumstances in which there is no significant risk that the qui tam lawsuit will harm the client, a lawyer-relator would not be using the information “to the disadvantage of the client,” unless deprivation of a business opportunity counts as such a disadvantage. Some
jurisdictions, however, more broadly prohibit use of a client’s confidential information to either the client’s disadvantage or to the advantage of the lawyer or a third person, except with the client’s informed consent. Moreover, even in jurisdictions that follow the narrower Model Rule approach, the lawyer’s common-law fiduciary duty not to engage in self-dealing requires the lawyer to account for any profits acquired through the use of client information unless the client has consented to that use. Given that the use of the information will not be to the detriment of the client, and that the lawyer and client can work together to serve as qui tam relators, it is possible that a client will consent both to any conflict of interest and to the lawyer’s use of the client’s information, thereby permitting the lawyer to serve as an FCA whistleblower against a third party defendant.

2. Loyalty Obligations to Former Clients

In FLPA I, a federal district court held that the company’s former counsel was barred from serving as a qui tam relator, regardless of whether he had impermissibly disclosed confidential client information. This was because his role in the lawsuit implicated him in an impermissible conflict of interest with his former client, in violation of the New York Lawyer’s Code of Professional Responsibility. The then-applicable New York rule, which was essentially the same as Model Rule 1.9(a), provided: “[A] lawyer who has represented a client in a matter shall not, without the consent of the former client after full disclosure . . . thereafter represent another person in the same or a substantially related matter in which that person’s interests are materially adverse to the interests of the former client.”

FLPA had argued that the rule did not apply because neither FLPA nor the former General Counsel was acting as a lawyer representing a client in the qui tam lawsuit. The district court, however, agreed with the defendant company that it was sufficient for purposes of the conflicts rule that the

---

196 MODEL CODE OF PROF’L RESPONSIBILITY DR 4-101(B). The District of Columbia, Michigan, Texas, and Virginia have similar requirements. See D.C. RULES OF PROF’L CONDUCT r. 1.6(a)(2)–(3); MICH. RULES OF PROF’L CONDUCT r. 1.6(b)(1)–(3); N.Y. RULES OF PROF’L CONDUCT r. 1.6(a); VA. RULES OF PROF’L CONDUCT r. 1.6(a) (VA. STATE BAR ASS’N 2015).

197 See RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 60(2).


199 See id.; see also FLPA II, 734 F.3d at 167–68. The Second Circuit affirmed the district court’s decision that the lawyer-relator had violated state confidentiality rules, but because it did not need to do so, it did not consider whether the lawyer had also violated the former-client conflict rule. FLPA II, 734 F.3d at 165.

former General Counsel, as an FLPA partner, was suing as a representative of the United States and not in his personal capacity.201

In so holding, the court accepted two separate arguments made by the defendant. The first argument relied on a literal reading of the rule, which, like the Model Rule, applied to a lawyer “who represents another person” and did not expressly require that the representation be that of a lawyer representing a client. In this regard, the court found that in a qui tam lawsuit, the qui tam relator “represents” the United States, relying both on the language of the FCA and on prior case law holding that a qui tam relator sues as a representative of the United States. The United States “remains the real party in interest.”202

The second argument accepted by the court was that the rule should not be interpreted in such a manner as to permit a former lawyer “to do directly, as a [representative] party, what he cannot do indirectly, as counsel.”203 Here the court cited decisions holding that, although a lawyer may sue a former client to vindicate personal rights, a lawyer may not bring a shareholder derivative action or serve as a class representative in a lawsuit against a former client (in the same or a substantially related matter), because serving as a representative plaintiff “implicates considerations distinct from affording an attorney the opportunity to vindicate rights personal to him.”204 These decisions, according to the court, reflect the view taken by the defendant’s expert: that to refuse to apply the former-client conflicts rule to lawyers acting as a representative party would “destroy one of the policies behind the [rule]—to encourage clients to trust and be candid with counsel,”205 which the court viewed as “a cornerstone of the lawyer-client relationship.”206

As the relator correctly noted, there was some precedent, contrary to the decisions cited by the court, holding that a lawyer who serves as a party representative does not violate the former-client conflicts rule. For example, in 2008, in Schaefer v. General Electric Co., a federal district court in Con-
necticut held that a former in-house lawyer could bring a Title VII sex discrimination claim both in her individual capacity and on behalf of a class of similarly situated female executive employees and attorneys.207 The court summarily refused to apply the former-client conflicts rule because the plaintiff was not representing a client in the lawsuit.208 Instead, the court considered whether, under the particular circumstances of that case, it was likely that the lawyer would reveal more client confidences than were necessary to establish her individual claim.209 Given that her individual allegations were based on non-confidential and non-privileged information, and her class allegations would be proved through the use of statistical and personal information, the court concluded that disqualifying the lawyer from serving as a class representative was unwarranted.210

Contrary to Schaefer, however, most courts have held that lawyers may not serve as a representative party in a shareholder derivative or class action lawsuit if they could not have represented a new client in the lawsuit.211 Nevertheless, regardless of whether these cases were correctly decided, they are arguably distinguishable from situations in which the former lawyer seeks to serve as a qui tam relator.

Many of the earlier shareholder derivative and class action cases were decided at a time when the disciplinary rules did not have a specific former-client conflicts rule.212 Instead, courts developed the “substantial relation-

---


208 Schaefer, 2008 U.S. Dist. LEXIS 5552, at *36. The court did not address the fact that the rule does not expressly require that type of representation. See id.

209 Id.

210 Id. at *17–36.

211 In addition to the cases cited in the district court’s opinion in FLPA I, 2011 U.S. Dist. LEXIS 37014, at *27–30, there are a significant number of federal circuit court opinions directly on point. See generally Doe v. A Corp. (A Corp. I), 709 F.2d 1043 (5th Cir. 1983) (barring former in-house counsel for the defendant corporation from prosecuting a class action but holding that he could prosecute the action on his own behalf); Richardson v. Hamilton Int’l Corp., 469 F.2d 1382 (3d Cir. 1972) (disqualifying former counsel for the defendant company from maintaining a class action and shareholder derivative suit against the company and certain officers and directors); Hall v. A Corp., 453 F.2d 1375 (2d Cir. 1972) (disqualifying the defendant’s former lawyer from serving as a class representative).

212 The federal circuit court cases cited above, supra note 211, were decided prior to or in the same year that the ABA Model Rules of Professional Conduct were first adopted in 1983; thus none of them relied on Model Rule 1.9(a) or a state counterpart. For a discussion of the absence of a former-client conflict rule prior to the 1983 adoption of the Model Rules, see infra notes 215–217 and accompanying text.
ship” test (subsequently codified in Model Rule 1.9(a) and its state counterparts) as part of a common law of lawyer disqualification designed to prevent lawyers from impermissibly disclosing confidential client information of their former clients.\footnote{213} With respect to lawyers bringing shareholder derivative actions, there was presumably no confidentiality exception that would have permitted them to disclose confidential information in pursuit of the lawsuit; therefore, it was probably necessary to disqualify these lawyers from acting as party-representatives in order to prevent them from violating their confidentiality obligations.\footnote{214} As for lawyers attempting to serve as class action representatives, prior to the adoption of the Model Rules in 1983, there was no confidentiality exception permitting disclosure in pursuit of the lawyer’s personal claim against a former client.\footnote{215} Therefore, disqualification of the lawyer was similarly necessary to protect the defendant’s confidential information. At least one court permitted a lawyer to pursue a personal claim against a former client, because there was no “social interest in allowing [the former client] to deny [the lawyer] . . . rights or . . . benefits if they are legally due him,” but nevertheless disqualified the lawyer from serving as a class representative because of the risk that he would disclose more information than was necessary in pursuit of his own claim.\footnote{216} Prior to the \textit{FLPA} decisions, no court had decided the disqualification question in the context of a lawyer representative such as a qui tam relator who was authorized to disclose a potentially wide range of client information under confidentiality exceptions designed to prevent or rectify former-client crimes or frauds.\footnote{217}

\footnote{213} See 1 HAZARD & HODES, supra note 152, § 14.07; CHARLES WOLFRAM, MODERN LEGAL ETHICS 363 (1983).

\footnote{214} See, e.g., Richardson, 469 F.2d at 1385–86 (disqualifying a lawyer from maintaining both a class action and a shareholder derivative suit because he may have acquired information in his prior employment that would be used in the present action).

\footnote{215} See supra notes 11–25 and accompanying text; see also \textit{A Corp. I}, 709 F.2d at 1047–48 (barring the plaintiff from maintaining a class action but saying he could prosecute an action on his own behalf).

\footnote{216} See, e.g., \textit{A. Corp. I}, 709 F.2d at 1050. In contrast, the district court in \textit{Schaefer} found that the lawyer’s individual allegations were based on non-confidential and non-privileged information, and that her class action allegations would be proved through the use of statistical and personal information, thereby entailing no significant risk of abuse of the defendant company’s confidential information. See \textit{Schaefer}, 2008 U.S. Dist. LEXIS 5552, at *36; see also Bakeman, 2006 Del. Ch. LEXIS 180, at *34–35.

\footnote{217} In \textit{Bury}, 2002 Cal. App. Unpub. LEXIS 1035, the court disqualified a lawyer from serving as a qui tam relator against his former employer on both confidentiality and conflict of interest grounds; however, unlike most other jurisdictions, California recognizes no exceptions to confidentiality to prevent or rectify merely economic harm. See \textsc{Cal. Bus. & Prof. Code} § 6068(e)(2) (West 2004) (providing that the sole exception is to prevent a criminal act likely to result in death or substantial bodily harm).
In our view, it was not inevitable that the FLPA II court would decide that New York’s Rule 1.9(a) prohibits a lawyer from serving as an FCA relator against a former client. Prior decisions disqualifying lawyers bringing shareholder derivative and class action lawsuits could have been viewed not as applying a disciplinary rule, but rather as applying a common law of lawyer disqualification,218 in which the courts were arguably assessing, on a case-by-case basis, the need to disqualify the lawyer to prevent likely breaches of the lawyer’s confidentiality obligation.219

In support of this argument, the relator could have pointed out that, although the Model Rule on which the New York rule was based was not expressly limited to representation as a lawyer, this narrower interpretation may be what the rule drafters had in mind. Thus, the very first sentence in comment 1 of Model Rule 1.9 states that “[a]fter termination of a client-lawyer relationship, a lawyer has certain continuing duties with respect to confidentiality and conflicts of interest and thus may not represent another client except in conformity with this Rule.”220 Indeed, the remainder of comment 1 of Model Rule 1.9 apparently assumes that the current representation involves a lawyer-client relationship, and most of the case law interpreting this and similar rules involves lawyers representing current clients in matters adverse to a former client.221 Moreover, the Restatement (Third) of the Law Governing Lawyers specifically provides that, without the consent of both clients, “a lawyer who has represented a client in a matter may not thereafter represent another client in the same or a substantially related matter in which the interests of the former client are materially adverse.”222

---

218 See generally Keith Swisher, The Practice and Theory of Lawyer Disqualification, 27 Geo. J. Legal Ethics 71 (2014) (analyzing lawyer disqualification cases involving conflict of interest, appearance of impropriety, and misconduct). Even after the ABA promulgated a model former-client conflicts disciplinary rule in 1983, courts did not automatically apply that rule in determining whether to disqualify a lawyer; rather, they further developed the common law of disqualification, taking into account some considerations that do not apply in the disciplinary context. See id. at 80.

219 See supra notes 207–210 and accompanying text (referencing Schaefer and Bakeman, cases in which each court conducted an individual assessment of the likelihood of improper disclosure of confidential information).

220 MODEL RULES OF PROF’L CONDUCT r. 1.9(a) cmt. 1 (emphasis added). The New York Code referenced in FLPA I did not have comments; even the current New York Rules of Professional Conduct, which are based on the ABA Model Rules, do not have official comments. See Roy D. Simon, Simon’s New York Rules of Professional Conduct Annotated 2 (2014) (noting that only black-letter rules have been adopted by courts; comments are adopted only by the New York State Bar Association and are therefore explanatory, not binding).

221 See MODEL RULES OF PROF’L CONDUCT r. 1.9(a) cmt. 1; Bennett et al., supra note 157, § 1.9 (citing several cases in which a lawyer attempted to represent a new client in a matter adverse to a former client).

222 RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 132 (emphasis added).
None of the comments nor the reporter’s notes to this Restatement provision suggest any awareness that the rule as stated differs from Model Rule 1.9(a).

As courts sometimes acknowledge, rules of professional conduct are not statutes enacted by a legislature. As a result, courts have often felt more at liberty to interpret these rules (unlike statutes) in a way that advances the court’s own understanding of the purposes underlying the rules, taking into account other law that is or may be applicable to lawyers. Thus, the FLPA I court concluded that the former-client conflict rule should be interpreted to apply to any instance in which a lawyer acts in a representative capacity, on the ground that lawyers should not be able to do directly as parties what they cannot do indirectly in representing a client. On the contrary, the relator in FLPA I had argued that the spirit of the rules as a whole would not be violated if lawyers were permitted to sue as qui tam relators, so long as they do not disclose more information than is necessary to prevent or rectify the former client’s crime or fraud. This argument might be rephrased to urge that respect for the federal interests underlying the FCA requires courts to follow the admonition of the federal district court in Doe that “to the extent that state law permits a disclosure of client confidences, such as to prevent a future or ongoing crime or fraud, then the attorney’s use of the qui tam mechanism to expose that fraud should be encouraged, not deterred.”

Which view better reflects the differing policies underlying both the rules of professional conduct and the False Claims Act? To answer this question, it is first necessary to understand the policies underlying the former-client conflicts rule under Model Rule 1.9(a) and its state counterparts, as well as the common law of lawyer disqualification, in core cases involving the representation of a new client adverse to a former client in the same or a substantially related matter.

---

223 See id. cmts. a–j & reporter’s notes.
224 FLPA I, 2011 U.S. Dist. LEXIS 37014, at *27–30 (finding that the New York Code equivalent of Model Rule 1.9(a) does not permit a former attorney “to do directly, as a party, what he could not do indirectly, as counsel”).
225 See Brief and Special Appendix for Plaintiff-Appellant, supra note 207, at 26–31.
227 We do not mean here to address the question of federal preemption of state disciplinary codes, which we address in Part III. Rather, what we are suggesting here is that when courts are interpreting a disciplinary code provision, they may and should consider the context in which the interpretive issue is raised.
Although there are some differences in the application of the former-client conflicts rule, it is generally agreed that, in these core cases, lawyers may not avoid disqualification by attempting to establish either that they have no confidential information of the former client that would be of use to the new client or that, even if they have such information, they will not impermissibly disclose it. Rather, courts use the “substantial relationship” test to identify situations in which it is likely that the lawyer had access to confidential information of the former client that would be of use to the new client.\textsuperscript{228} In such situations, courts presume that there is a sufficient threat to the impermissible use of confidential client information of the former client that the lawyer must be prevented from undertaking the new representation.\textsuperscript{229} This presumption is irrebuttable, because to allow it to be rebutted would force the client to reveal the very information that it wishes to keep confidential and would put the parties in the awkward position of debating whether this particular lawyer is capable of adhering to his or her confidentiality obligations.\textsuperscript{230}

The rationale for presumptive disqualification appears to be the need for prophylactic rules that prevent a lawyer from impermissibly using or disclosing a former client’s confidential information. Therefore, would-be lawyer-relators will argue that when they are permitted to use or disclose a former client’s confidential information in order to prevent or rectify the former client’s crime or fraud, there is no reason to prevent them from acting adversely to the former client in a substantially related matter.\textsuperscript{231} As a result, the former-client conflicts rule should not be interpreted to apply to such qui tam relators (or, regardless of how the disciplinary rule is interpreted, these lawyers should not be disqualified from serving as qui tam relators).

There are two possible responses to this argument. The first response is that prophylactic disqualification of the lawyer might still be necessary to ensure that the lawyer uses or discloses only that information that is necessary to prevent or rectify the former client’s crime or fraud. In other words, although the lawyer might be free to disclose information to the government

\textsuperscript{228} See, e.g., 1 HAZARD & HODES, supra note 152, § 14.07.

\textsuperscript{229} See id.

\textsuperscript{230} See id. There is an additional concern that a lawyer will be restricted in the representation of a current client. But this concern is relevant to determining the lawyer’s obligations under Model Rule 1.7, not under Rule 1.9(a), the latter being designed to protect the interests of the former client. See MODEL RULES OF PROF’L CONDUCT r. 1.7, 1.9(a).

\textsuperscript{231} Cf. Oasis W. Realty, LLC v. Goldman, 250 P.3d 1115 (Cal. 2011) (sustaining a breach of fiduciary duty claim against a lawyer who in his personal capacity publicly protested a development permit that he himself had formerly obtained on behalf of the client, and declining to limit precedents to situations involving the successive representation of clients).
(without attempting to serve as a qui tam relator), permitting the lawyer to serve as a qui tam relator provides too great an incentive for the lawyer to use or disclose more information than would be permitted under the applicable exception to the confidentiality rule.

This response is weak because it insufficiently acknowledges the existence of important interests other than those of the former client. Even in core cases involving representation of a current client adverse to a former client in a substantially related matter, it is understood that the substantial relationship test is underinclusive and will not identify all cases in which the lawyer has confidential information from the former client that could be useful to the new client in pursuing the current litigation. Nevertheless, neither Model Rule 1.9(a) nor the common law of disqualification prohibits lawyers from representing current clients adverse to their former clients in situations in which there is merely a remote possibility that the lawyer has obtained relevant information. Rather, prophylactic disqualification is limited to the same or a substantially related matter in order to avoid unduly burdening the current client’s interest in hiring counsel of choice, as well as the lawyer’s legitimate interest in taking on new clients and new matters in furtherance of the lawyer’s career. When the matters are not the same or substantially related, courts trust lawyers to adhere to their confidentiality obligations to their former clients by identifying situations in which confidential information is at risk and then either voluntarily declining the new matter or taking care to avoid impermissibly using or disclosing such information in the course of the current representation. Similarly, given that there may be much information that qui tam relators are permitted to disclose in prosecuting the qui tam lawsuit, courts should trust them not to disclose more information than is permitted under the applicable confidentiality exception: to do otherwise would insufficiently acknowledge the federal interest in having insiders with relevant information blow the whistle on companies that have defrauded the federal government. And if the lawyer does disclose information unnecessarily, then the lawyer is still subject to discipline for breaching the obligation of confidentiality.

A second response to arguments favoring a narrow interpretation of Model Rule 1.9(a) is perhaps more compelling. The substantial relationship test is a prophylactic standard aimed primarily at protecting the former cli-

---

232 See, e.g., 1 HAZARD & HODES, supra note 152, § 14.07 (stating that it is insufficient that the lawyer for a client in one matter “could have” obtained confidential information about the other).

233 See id.

234 Cf. id. (“It is unfair to the lawyer and to the lawyer’s new client for a former client to be able to paint an overly broad picture of the prior representation without fear of contradiction.”).
rent’s confidential information; however, in addressing subsequent adverse representation, some courts have focused not just on this confidentiality concern, but also on the lawyer’s duty of loyalty.235 Most courts, however, limit the independent role of loyalty under state versions of Rule 1.9(a) to several narrow situations; for example, those involving an attack on the lawyer’s own work for the former client and those involving former clients who were jointly represented, where there is typically no expectation of confidentiality between the clients.236 In both instances, the purpose of the prohibition on subsequent adverse representation is to prevent the lawyer from acting improperly in the earlier representation. For example, if the former-client conflicts rule did not apply to attacks on a lawyer’s own prior work product, then a lawyer might draft documents that are susceptible to a later challenge, thereby creating the potential for lucrative work on behalf of a new client. Even if such traps were laid inadvertently, clients should be assured that the lawyer will not be permitted to exploit them on behalf of a new client.237 Similarly, and perhaps even more compellingly, if a lawyer is permitted to proceed adversely to a former joint client, then the lawyer would have an incentive to impermissibly favor the other client during the period of the common representation, particularly if the common representation is about to end.238 In both situations, the purpose of prohibiting subsequent representation is to neutralize the risk of disloyalty during the earlier representation.

With respect to lawyers serving as qui tam relators, it may be necessary or desirable to prevent them from subsequently “representing” the government in matters substantially related to work they performed for their former client in order to ensure against disloyalty during the period that the lawyer is representing the company.239 For example, while employed by the

235 See Charles W. Wolfram, Symposium, Former-Client Conflicts, 10 GEO. J. LEGAL ETHICS 677, 691–712 (1997) (noting that “some courts and commentators also take the position that an analytically separate and important consideration in determining the scope of the former-client conflict rules is that of loyalty”; also criticizing the appeal to loyalty concerns except in several specific instances); see also Bury, 2002 Cal. App. Unpub. LEXIS 1035, at *11. In Bury, the court stated:

It is . . . an attorney’s duty to protect his or her client in every possible way. An attorney violates this duty of loyalty if he or she assumes a position adverse or antagonistic to a client without the client’s free and intelligent consent. The duties of confidentiality and of loyalty are owed to former as well as present clients.


237 Id. at 697–98.

238 Id. at 711.

239 See N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 13–14. Another rationale was expressed in a recent opinion by the New York County Lawyers’ Association’s Committee on Professional
A lawyer who has the option of leaving the company (either voluntarily or involuntarily) and then filing a qui tam lawsuit against the company could easily be tempted to act in a manner that would make that lawsuit more likely to succeed. Preventing this from occurring could justify either interpreting the former-client conflicts rule to cover subsequent “representation” by a qui tam relator or disqualifying the lawyer-relator regardless of whether the former-client conflicts rule expressly applies.

If Model Rule 1.9(a) precludes lawyers from filing a qui tam lawsuit against a former client, then other potential loyalty obligations are clearly moot. But there is a good chance that courts will ultimately decide that Rule 1.9(a) does not apply because qui tam relators do not “represent” the government in the sense intended under the rule. In that case, we should also consider whether the lawyer has other applicable loyalty obligations. As we discussed previously, lawyers have a duty under Model Rule 1.4 to keep their current clients informed of significant developments in the representation; however, the rule does not generally apply to former clients. The prohibition on self-dealing, however, continues even after a representation has ended. As with a current client, this common-law duty appears to broadly prohibit a lawyer from serving as a qui tam relator, even against a former client, unless federal law preempts the common-law duty.

Ethics with respect to whistleblower bounties under Dodd-Frank. After noting that a disclosure is not permitted under confidentiality exceptions to the New York Rules of Professional Conduct unless it is “reasonably necessary,” the Committee concluded that:

Undertaking this otherwise permissible disclosure in a manner that results in a bounty for the lawyer raises a significant risk that the attorney’s judgment in determining whether the disclosure is “reasonably necessary” will be adversely affected and presents a conflict of interest that is beyond what Rule 1.9 was intended to allow.

Interestingly, the opinion does not mention New York Rule 1.9(a), the former-client conflicts rule that was expressly addressed in the district court’s decision in FLPA I. It would be difficult to extend the rationale of FLPA I to whistleblower bounties under Dodd-Frank because there is no sense in which the Dodd-Frank whistleblower acts in a representative capacity, as is the case under the FCA. See infra notes 296–301 and accompanying text.

See, e.g., N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 10; Temkin & Moskovits, supra note 8, at 14.

See supra notes 26–123 and accompanying text.
II. SEC WHISTLEBLOWER AWARDS UNDER DODD-FRANK

A. Primer on Dodd-Frank Whistleblower Awards

Dodd-Frank requires the SEC to give financial awards to whistleblowers who provide it with information that results in successful enforcement actions. Under this statute, if a whistleblower gives the SEC “original information,” and the SEC brings an enforcement action that yields monetary sanctions greater than $1 million, then the SEC must pay the whistleblower between 10–30% of the sanction amount. The statute excludes five categories of individuals from these whistleblower awards: (1) employees of certain government agencies; (2) employees of self-regulatory organizations; (3) anyone convicted of a crime related to the enforcement action; (4) anyone who knowingly and willfully provides false information; and (5) anyone who learned the information through the performance of a statutorily required audit.

The SEC’s Dodd-Frank whistleblower regulations identify the government agencies and self-regulatory organizations whose employees are excluded from the program. These regulations also identify additional categories of individuals and information that are excluded from the whistleblower award program, including SEC employees’ family members, anyone who provides information in response to a subpoena, and anyone who obtained the information by means that a court has determined were unlawful. The regulations permit a person involved with a corporation’s compliance function to qualify for a whistleblower award if he or she reported it internally at least 120 days prior to reporting it to the SEC.

The SEC’s whistleblower award process has several distinct stages. First, the whistleblower must submit information about the alleged violation.

---


244 Id. § 78u-6.


246 Id. § 240.21F-8(c)(3)–(7).

247 Id. § 240.21F-4(b)(4)(v)(c).
by filling out a specific form. Second, the SEC investigates the allegation. If that investigation leads to an enforcement action and if that enforcement action results in sanctions over $1 million, then the whistleblower may apply for an award. Every few months, the SEC publishes a list of all of its enforcement actions that have resulted in sanctions over $1 million (including those resulting from whistleblower tips and those resulting from other sources). At that point, a whistleblower has ninety days to claim a whistleblower award by filling out another form, an Application for Award. The SEC determines whether the whistleblower is eligible for an award, and, if eligible, the amount of the award. If the SEC denies an award or grants an award less than the statutory ten percent minimum, then the whistleblower can appeal that determination to a federal circuit court.

The program is still in its infancy because it takes time for a whistleblower case to percolate through the SEC’s investigative and administrative processes. In the few years since the SEC issued regulations implementing the statute, the agency has received about 3000 whistleblower tips per year and has paid more than $54 million to twenty-two whistleblowers, including one award for more than $30 million.

Unlike the FCA, which requires whistleblowers to publicly identify themselves and thus risk retaliation, the SEC’s Dodd-Frank program allows whistleblowers to keep their identity secret. The statute permits whistleblowers to submit tips anonymously as long as they do so through an attor-

---

248 Id. § 240.21F-9(a). The form asks questions about the individual’s potential eligibility for a whistleblower award, including whether the individual is or was the entity’s counsel, and whether any of the information “was obtained from an attorney or in a communication where an attorney was present.” Form TCR: Tip, Complaint or Referral, SECs. & EXCH. COMM’N (2011), https://www.sec.gov/about/forms/formtcrr.pdf [http://perma.cc/6EXG-E7RR] (see questions 5a and 10).
249 17 C.F.R. § 240.21F-10(d).
250 Id. § 240.21F-10(a).
251 Id.
252 Id.; see also id. § 240.21F-10(b) (referring to “Form WB–APP, Application for Award for Original Information Provided Pursuant to Section 21F of the Securities Exchange Act of 1934”).
253 Id. § 240.21F-10(d).
254 Id. § 240.21F-13(a).
256 See 15 U.S.C. § 78u-6(h)(2)(A) (stating that “the Commission shall not disclose any information, including information provided by a whistleblower to the Commission, which could reasonably be expected to reveal the identity of a whistleblower” except in several limited circumstances).
To claim an award, a whistleblower needs to reveal his or her identity to the SEC, but the statute prohibits the SEC from "disclos[ing] any information . . . which could reasonably be expected to reveal the identity of a whistleblower." When the SEC announces that it has made a whistleblower award, it describes the facts leading to the award in such generic terms that it is usually impossible to identify the company involved (let alone the whistleblower). Thus, in the limited history of the SEC’s Dodd-Frank whistleblower program, the SEC has taken seriously its statutory mandate to protect the identity of whistleblowers. Because of the secrecy surrounding the SEC’s whistleblower program, we simply do not know whether any of its awards have gone to lawyers.

If a company learned that its current or former lawyer disclosed its alleged wrongdoing to the government and is seeking a financial award for doing so, the company could be expected to object on grounds of confidentiality or loyalty. But under Dodd-Frank, a company may never find out. The entire process takes place in secret within an administrative agency. Although a whistleblower could seek judicial review if the SEC denies an award entirely or grants an award that is less than ten percent of the sanctions, it is not clear that a client company will ever be in a position to seek judicial review of an SEC award to a lawyer-whistleblower (or any other kind of whistleblower, for that matter). Nevertheless, we expect that the SEC will abide by its own mandate not to grant awards where a lawyer’s disclosures go beyond that permitted by state confidentiality rules or the SOX regulation.

257 Id. § 78u-6(d)(2)(A); 17 C.F.R. § 240.21F-7(b) (indicating that the attorney must know the whistleblower’s identity).
258 15 U.S.C. § 78u-6(d)(2)(B); 17 C.F.R. § 240.21F-7(b)(3).
B. Lawyers’ Confidentiality Obligations and Dodd-Frank Whistleblower Awards

The Dodd-Frank statute indicates that whistleblower awards are available to individuals who provide the SEC with “original information” that leads to a successful enforcement action. Although the statute excludes several classes of individuals from being eligible for these awards, it does not address whether lawyers are eligible or whether lawyers seeking such awards may reveal otherwise confidential information. The SEC regulation defining “original information” does exclude certain types of information from becoming the basis for an award, including information obtained from communications subject to the attorney-client privilege or obtained in connection with legal representation. This regulation would ordinarily prevent lawyers from obtaining whistleblower awards for disclosing their clients’ securities violations. But what the regulation takes away with one hand, it partially gives back with another: information that is privileged or obtained in connection with legal representation may nonetheless form the basis of an award if a lawyer would be permitted to disclose that information under applicable state attorney conduct rules, under the SEC’s earlier SOX regulation, or otherwise. For lawyers who are potential whistleblowers, this is perhaps the most important feature of the SEC’s Dodd-Frank whistleblower program: the fact that it incorporates the SEC’s SOX regulation, which creates new confidentiality exceptions for lawyers practicing before the SEC.

In issuing its Dodd-Frank whistleblower regulation, the SEC asserted that the regulation “strikes the right balance because” it is “consistent with the public policy judgments” in its earlier SOX regulation “as to when the benefits of permitting disclosure are justified notwithstanding any potential harm to the attorney-client relationship.” Because the SEC’s Dodd-Frank whistleblower regulation incorporates the earlier SOX regulation, it is necessary to examine the SOX regulation’s confidentiality exceptions in some detail.

263 Id. § 78u-6(c)(2). The statute excludes several categories of individuals from receiving whistleblower awards, including employees of the Justice Department and law enforcement organizations. Id.
264 17 C.F.R. § 240.21F-4(b)(4)(i).
265 Id. § 240.21F-4(b)(4)(ii).
266 Id. § 240.21F-4(b)(4)(i)-(ii).
Section 307 of the SOX Act required the SEC to “issue rules . . . setting forth minimum standards of professional conduct for attorneys appearing and practicing before the Commission.”269 The statute specified that the SEC’s regulation must “includ[e] a rule” requiring that lawyers with “evidence of a material violation of securities law or breach of fiduciary duty” must engage in internal whistleblowing, ensuring that a corporation’s leadership would be aware of that evidence.270 In response to that statute, the SEC issued a regulation requiring such a lawyer to engage in internal whistleblowing if the lawyer is aware of “credible evidence” that it is “reasonably likely that a material violation” of securities law or a material breach of a fiduciary duty has occurred, is ongoing, or is about to occur.271 This mandate for internal whistleblowing is similar to that found in ABA Model Rule 1.13(b), but it has a lower evidentiary trigger: credible evidence of a reasonably likely violation, rather than 1.13(b)’s knowledge of a violation.272 The regulation also included several additional provisions that were “not explicitly required by Section 307, but which the Commission believe[d] [were] important components of an effective ‘up the ladder’ reporting system.”273 One of those additional provisions created new confidentiality exceptions for lawyers practicing before the SEC, allowing them to disclose information to the SEC “to the extent the attorney reasonably believes necessary” to “prevent the issuer from committing a material violation” or to “rectify the consequences of” the issuer’s past “material violation . . . that caused, or may cause, substantial injury to the financial interest or property of the issuer or investors in the furtherance of which the attorney’s services were used.”274

270 Id.
271 17 C.F.R. § 205.2(e) (defining “evidence of a material violation”); id. § 205.2(i) (defining “material violation”); id. § 205.3(b) (providing the mandate for internal whistleblowing). The proposed rule also included several other provisions that the SEC omitted from its final rule, including provisions requiring lawyers to notify the SEC of a client’s material violations. See, e.g., Implementation of Standards of Professional Conduct for Attorneys, Securities Act Release No. 8150, Exchange Act Release No. 46,868, Investment Company Act Release No. 25,829, 67 Fed. Reg. 71,670, 71,705 (proposed Dec. 2, 2002) (proposing 17 C.F.R. § 205.3(b)(3), (d)).
272 See MODEL RULES OF PROF’L CONDUCT r. 1.13(b) (AM. BAR ASS’N 2015).
274 17 C.F.R. § 205.3(d)(2)(iii). It is this provision—allowing lawyers to disclose information to rectify a client’s fraud—that is likely to be most relevant to lawyers seeking Dodd-Frank whistleblower awards. The evidentiary trigger permitting external whistleblowing is higher than that mandating internal whistleblowing. External whistleblowing is permissible only if the lawyer “reasonably believes” that the disclosure is necessary in order to prevent or rectify a violation—a standard similar to that found in Model Rule 1.6(b)(2) and (b)(3). Id.; see also MODEL RULES OF PROF’L CON-
The SOX regulation’s confidentiality exceptions are, at the highest level of generality, similar to the fraud-related exceptions found in the Model Rules and most states’ rules. The confidentiality exceptions in both are triggered if the lawyer “reasonably believes” that disclosure is “necessary” in order to prevent or rectify certain violations. But at a more granular level, the differences are significant and the SOX exceptions are generally broader in scope. Instead of being limited to crimes and frauds, the SOX exceptions reach “material violations” of federal or state securities law and material breaches of fiduciary duty. Unlike Model Rule 1.6, the SOX regulation permits lawyers to disclose in order to prevent a violation even if the lawyer’s services were not used in that violation. It permits lawyers to disclose in order to rectify a past violation or breach when their services were used in furtherance of the violation or breach, even if it is the issuer (rather than a third party) that is injured by the violation.

The SEC’s SOX regulation does not apply to all lawyers. Instead it applies to (and can be invoked by) lawyers “[a]ppearing and practicing before” the SEC. This phrase—“[a]ppearing and practicing”—reaches not just lawyers who “appear” in the traditional sense of “[r]epresenting an issuer in a Commission administrative proceeding,” but also lawyers who provide advice about whether information must be submitted to the SEC or advice about the adequacy of any document submitted, as well as any lawyer “[t]ransacting any business with the SEC, including communications in any form.” It has been suggested that when a whistleblowing lawyer makes a disclosure to the SEC in order to later qualify for an award, that
disclosure constitutes a “communication” that qualifies the lawyer as “appearing and practicing” before the SEC, making broad SOX confidentiality exceptions applicable. But a lawyer's communication with the SEC constitutes “appearing and practicing” only if it occurs “in the context of providing legal services to an issuer with whom the attorney has an attorney-client relationship.”286 If the disclosure is for the purpose of qualifying for a whistleblowing award, it would be outside “the context of providing legal services to an issuer,” and therefore this bootstrapping approach would not be available.

In the context of FCA-related disclosures, a lawyer’s ability to disclose information largely depends on whether the applicable state confidentiality rule permits disclosure to rectify, mitigate, or prevent client frauds and crimes, or prohibits such disclosures entirely. In the context of Dodd-Frank, any lawyer who appears and practices before the SEC can avail him or herself of the expanded confidentiality exceptions found in the SOX regulation—and therefore may inform the SEC of material securities law violations and material breaches of fiduciary duty even if the applicable state confidentiality rule would prohibit such a disclosure. But the key question under Dodd-Frank is whether the lawyer who wishes to use the SOX exceptions is actually subject to the SOX regulation.

C. Lawyers’ Loyalty Obligations and Dodd-Frank Whistleblower Awards

1. Current Clients

No court has yet addressed the issue of the ethical propriety of lawyers seeking whistleblower bounties under the Dodd-Frank legislation. Various commentators287 and at least one bar ethics committee,288 however, have addressed the applicability of state confidentiality and/or conflicts rules under the SEC regulations implementing the Dodd-Frank whistleblower bounty provisions. Regarding conflicts of interest, most of the attention has been focused on conflicts with current clients. This is probably because, unlike the FCA, Dodd-Frank permits whistleblowers to remain anonymous, even after the receipt of an award, thereby making it more likely that a lawyer will blow the whistle on a current client.289

286 Id. § 205.2(a)(2)(i).
287 See, e.g., Pacella, supra note 7, at 1040–45; Temkin & Moskovits, supra note 8, at 23; Green & Thomas, supra note 144; LATHAM & WATKINS, supra note 90, at 6–15.
288 See generally N.Y. Cnty. Lawyers’ Ass’n, supra note 8 (discussing whether a lawyer who is subject to the New York Rules of Professional Conduct may seek a Dodd-Frank whistleblower award).
289 Under Dodd-Frank, the SEC will ultimately need to learn the whistleblower’s identity for an award to be made, but the SEC has a statutory obligation to keep that information confidential.
Those who have addressed the issue agree, as do we, that just like an FCA lawyer-relator, a Dodd-Frank whistleblower-lawyer who is currently representing the company on the subject of the disclosure will likely violate Model Rule 1.7 with respect to “material limitation” conflicts.\(^{290}\) This is because there is a significant risk that the lawyer’s representation will be limited by the lawyer’s personal pecuniary interest in obtaining the whistleblower award, which is likely to compromise the lawyer’s ability to objectively consider and advise the company concerning such questions as whether the company is violating the law, whether the legal violation poses a threat to the company, and whether suspected wrongdoing should be reported to a higher level in the company.\(^{291}\) Moreover, such a conflict will be non-consentable if, given the amount of money at stake, the lawyer cannot reasonably believe that he or she can provide diligent and competent representation.\(^{292}\) More important, as in FCA cases, it is unlikely that a client would consent to permitting its lawyer to be a Dodd-Frank whistleblower.

Other aspects of the lawyer’s duty of loyalty under Dodd-Frank have not been as thoroughly discussed, including whether a lawyer may act adversely to a current client by disclosing client information that is not the subject of the current representation. Some commentators have assumed that continued representation of the client on an unrelated matter would be ethically permissible,\(^{293}\) without addressing whether lawyers are permitted

---

See 15 U.S.C. § 78u-6(h)(2)(A); 17 C.F.R. § 240.21F-7(a). Under the FCA, the filing of the complaint is done under seal, but the lawyer-relator’s name becomes public once the initial seal is lifted. See 31 U.S.C. § 3730(b) (2012).

\(^{290}\) See MODEL RULES OF PROF’L CONDUCT r. 1.7; Pacella, supra note 7, at 1048; Temkin & Moskovits, supra note 8, at 22–23; Green & Thomas, supra note 144.

\(^{291}\) Temkin & Moskovits, supra note 8, at 22–23; see also N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 11 (noting that “potential bounties range from $100,000 to literally millions of dollars in large cases”). According to the New York County Lawyers’ Association (“NYCLA”) opinion, an anticipated whistleblower bounty in excess of $100,000 “presumptively gives rise to a conflict of interest between the lawyer’s personal interest and that of the client.” N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 11. The Committee believes that a conflict of interest arises in “the overwhelming majority of cases.” Id. The Committee does not address in what circumstances such a personal interest would not give rise to a conflict, other than to suggest that if the lawyer is required to make a disclosure to someone, then “financial incentive could be less of a factor in determining the existence of a conflict with the lawyer’s personal interest.” Id. New York Rule 3.3(b) provides one example of such a required disclosure: when the lawyer “knows that a person intends to engage, is engaging or has engaged in criminal or fraudulent conduct related to [a proceeding before a tribunal].” N.Y. RULES OF PROF’L CONDUCT r. 3.3(b) (N.Y. STATE BAR ASS’N 2015).

\(^{292}\) See N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 11. The NYCLA opinion cautions that “[i]n some circumstances the whistleblower-bounty conflict may be unwaiveable,” but we believe it is more accurate to say that this will typically be the case. See id. at 12 (emphasis added).

\(^{293}\) See, e.g., Green & Thomas, supra note 144 (stating that a lawyer-whistleblower may continue representing a client after submitting information to the SEC unless the representation would
to act adversely to a current client, including whether fiduciary law requires them to avoid self-dealing even in unrelated information and to account for any profits received. Nor has there been any discussion of whether a lawyer must inform a client that the lawyer has disclosed information concerning the client’s wrongdoing to the SEC, or whether a lawyer may blow the whistle on a non-client third party without first seeking the client’s consent to take advantage of a financial opportunity that may belong to the client. As to all of these issues, we believe that the loyalty analysis under Dodd-Frank is substantially similar to the analysis under the FCA, with the result that, although the disciplinary rules do not clearly prohibit lawyer conduct that is directly adverse to a current client on an unrelated matter, or seeking a financial reward for disclosing information adverse to a non-client third party, fiduciary law continues to broadly prevent the lawyer from receiving a significant financial reward at the client’s expense.

2. Former Clients

There is little discussion of former-client conflicts under Dodd-Frank because, unlike the FCA, a lawyer-whistleblower does not act in any form of an arguably representative capacity. As a result, seeking a Dodd-Frank whistleblower award presumably would not violate Model Rule 1.9(a) or its state equivalents.

Nevertheless, a recent opinion of the New York County Lawyers’ Association’s Committee on Professional Ethics concluded that, even when a lawyer is permitted to disclose information under New York Rule 1.9(c),

[U]ndertaking this otherwise permissible disclosure in a manner that results in a bounty for the lawyer raises a significant risk that the attorney’s judgment in determining whether the disclosure is ‘reasonably necessary’ will be adversely affected and presents a conflict of interest that is beyond what Rule 1.9 was intended to allow.

be affected by the attorney’s status as a whistleblower; for example, if the lawyer was requested to advise the client how to respond to an SEC inquiry prompted by the lawyer’s disclosure).

294 One difference is that, although Dodd-Frank whistleblowers are taking action directly adverse to their current clients, they are not publicly suing them, as are FCA relators. Whether this difference is a significant one is less clear.

295 See, e.g., RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 60(2) (AM. LAW INST. 2000) (stating that a lawyer has a fiduciary duty to report profits made using client information unless the client has given consent for such use); see also supra note 197 and accompanying text.

296 See MODEL RULES OF PROF’L CONDUCT r. 1.9(a).

297 N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 14.
We disagree that New York Rule 1.9, which is based on Model Rule 1.9, is properly interpreted to prohibit a lawyer from seeking a whistleblower award. What little authority the Committee cited is clearly distinguishable, and courts should be reluctant to discipline a lawyer when there is no particular rule that can be plausibly interpreted to prohibit the conduct in question. Moreover, given that whistleblowing under Dodd-Frank does not involve litigation, there is no possibility of court disqualification based, not on the disciplinary rules themselves, but rather on the common law of lawyer disqualification.

Even so, we believe that Dodd-Frank lawyer-whistleblowers do have a conflict of interest when they blow the whistle on a former client, even when they are ethically permitted to use or disclose that information under applicable rules of professional conduct. Our conclusion is based not on the disciplinary rules themselves, but rather on the existence of common-law fiduciary duties that are broader than the specific fiduciary duties codified in state disciplinary rules. An attorney’s fiduciary duty extends to former clients and includes the duty not to engage in self-dealing and to account for profits made with confidential information obtained during the lawyer-client relationship. It may be that such breaches will be difficult to detect, given the anonymity permitted under the Dodd-Frank regulations, but it is at least possible that an actual or potential breach will be detected.

---

298 The Committee cited Oasis West Realty, LLC v. Goldman, a 2011 California Supreme Court case, as “sustain[ing] a breach of fiduciary duty claim against a lawyer who was disloyal to a former client when he publicly protested a development permit that he himself had formerly obtained on behalf of the client, at considerable expense,” but the Goldman opinion focused on the lawyer’s adverse use of confidential client information rather than an impermissible disclosure. Id. at 13; see Oasis W. Realty, LLC v. Goldman, 250 P.3d 1115, 1122 (Cal. 2011). California had no confidentiality exception that would have permitted either the use or disclosure in that case. See Cal. Rules of Prof’l Conduct r. 3-100(B) (State Bar of Cal. 2015) (permitting the discretionary disclosure of confidential information only when necessary to prevent a criminal act that is likely to result in substantial bodily injury or death). The Committee cited another case, Birnbaum v. Birnbaum, decided by the New York Court of Appeals in 1989, for the broad proposition that a fiduciary’s “general duty of fidelity requires avoidance of situations in which personal interests conflict with the interests of those owed a fiduciary duty,” involved a nonlawyer fiduciary who engaged in self-dealing while he was a partner in a general partnership, and says nothing about the obligations of a former lawyer who is ethically permitted to use or disclose confidential information under applicable disciplinary rules. N.Y. Cnty. Lawyers’ Ass’n, supra note 8, at 14 n.30; see Birnbaum v. Birnbaum, 539 N.E.2d 574 (N.Y. 1989).

299 The NYCLA opinion appears to rely on New York Rule 1.9(c), which is based on Model Rule 1.9(c), but the Committee nowhere explains how the language of that provision should be interpreted to prohibit adverse action when the disclosure itself is permitted by a confidentiality exception. See generally N.Y. Cnty. Lawyers’ Ass’n, supra note 8.

300 See Restatement (Third) of Agency § 8.02 (Am. Law Inst. 2006); see also Tri-Growth Ctr. City, Ltd. v. Sildorf, Burdman, Duignan & Eisenberg, 265 Cal. Rptr. 330 (Ct. App. 1990).
ver, although disciplinary action is both unlikely and probably unjustified, it may be possible for the company to seek to enjoin the lawyer from either disclosing information to the SEC pursuant to the whistleblower bounty program \textsuperscript{301} or collecting a whistleblower bounty. The company could also sue the lawyer to recover any profits received or for any damages attributable to the breach. Finally, if the SEC is convinced that lawyers violate their state law fiduciary duties when they seek a whistleblower bounty based on information obtained in the representation, the SEC may become reluctant to make such awards or to encourage such lawyers to report to them. The SEC might even consider amending its rules to clarify that lawyers who violate their common-law fiduciary duties to their clients are ineligible to participate in the whistleblower bounty programs.

Of course, our conclusions with respect to the existence of an unethical conflict of interest for Dodd-Frank lawyer-whistleblowers, as well as for FCA lawyer-relators, assume that neither state rules of professional conduct nor state common-law fiduciary duties are preempted by federal law, which is the question to which we now briefly turn.

\section*{III. FEDERAL PREEMPTION OF STATE ETHICS LAW}

\textit{A. The False Claims Act}

The three federal FCA cases that address lawyer-relators’ ethical duties, \textit{United States ex rel. Doe v. X Corp. ("Doe")}, \textit{United States ex rel. Fair Laboratory Practices Associates v. Quest Diagnostics, Inc. ("FLPA II")}, and \textit{United States ex rel. Holmes v. Northrop Grumman}, assert that the FCA does not preempt state confidentiality rules. \textsuperscript{302} Certainly there is nothing in the text of the FCA statute that could be interpreted as express preemption

\textsuperscript{301} A court would not prohibit disclosure of information to the SEC where applicable confidentiality rules permit such disclosure, but could possibly prohibit disclosure in the manner required to trigger eligibility for a whistleblower bounty award. Such a prohibition would likely inhibit many lawyers from disclosing the information at all.

of state regulation of attorneys. As for implied preemption, all three courts summarily dismissed this possibility, citing U.S. Supreme Court cases declaring that in areas of traditional state regulation, the presumption against preemption cannot be overcome unless Congress has made such a purpose “clear and manifest.”

In *Doe*, the defendant urged the court to interpret the FCA to exclude lawyers from serving as qui tam relators against their own clients. The court found no basis in either the text or the legislative history of the statute to do so. As for the defendant’s argument that permitting lawyers to serve as plaintiffs against their own clients “would . . . encourage attorneys to flout their ethical obligations and use their clients’ confidential information to their own economic advantage,” the court responded that the FCA does not preempt state confidentiality rules, “does not authorize [a relator] to violate state law,” and does not “immunize a relator” who violates state law. The court further noted that permitting lawyers to “report their clients’ ongoing or planned fraudulent practices against the government” undeniable serves the “Congressional purpose underlying the *qui tam* provisions [which] is ‘to enhance the Government’s ability to recover losses sustained as a result of fraud against the Government.’” The *Doe* court concluded: “[T]o the extent that state law permits a disclosure of client confidences . . . then the attorney’s use of the *qui tam* mechanism to expose that fraud should be encouraged, not deterred.”

---

303 See, e.g., *Doe*, 862 F. Supp. at 1507 n.12 (“Congress expressed no intent in the Act to preempt state laws governing the attorney-client relationship.”).

304 See, e.g., id. (“[C]ourts must ‘start[] with the assumption that the historic police powers of the States [are] not to be superseded by . . . Federal Act unless that [is] the clear and manifest purpose of Congress.’” (latter alterations in original) (quoting *Santa Fe Elevator Corp.*, 331 U.S. 218, 230 (1947)); see also *FLPA II*, 734 F.3d at 163 (“In areas of traditional state regulation, we assume that a federal statute has not supplanted state law unless Congress has made such an intention clear and manifest.” (quoting *Bates v. Dow Agrosciences LLC*, 544 U.S. 431, 449 (2005)); *Doe*, 862 F. Supp. at 1507 (“Absent clear legislative intent to upset this settled practice, state law regarding the attorney-client relationship cannot be deemed preempted.”)).

305 *Doe*, 862 F. Supp. at 1507 (“X Corp. protests that [allowing lawyers to serve as relators] severely impinges upon the common law relationship between attorney and client and could not have been intended by Congress.”).

306 Id. at 1506, 1508 (“Congress, in plain language, carefully fashioned specific exclusions to the class of eligible relators, and those exclusions did not include attorneys. . . . [T]he *qui tam* statute does not exclude lawyers or members of any particular profession from being relators . . . .”).

307 Id. at 1507 (“[W]here an attorney’s disclosure of client confidences is prohibited by state law in a given circumstance, that attorney risks subjecting himself to corresponding state disciplinary proceedings should he attempt to make the disclosure in a *qui tam* suit.”).


309 Id. at 1507–08 (footnote omitted).
In FLPA II, the relator argued for a form of FCA preemption in order to avoid application of a state ethics code that prohibited disclosing confidential client information except when reasonably necessary to prevent future or ongoing crimes or frauds. Although the U.S. Court of Appeals for the Second Circuit quickly concluded that the FCA does not preempt state ethics rules, the court then acknowledged “that the central purpose of the [New York] Rules—to protect client confidences—can be ‘inconsistent with or antithetical to federal interests’... which under the FCA, are to ‘encourage private individuals who are aware of fraud being perpetrated against the government to bring such information forward.’” Citing a prior disciplinary case, the court further explained that “[i]n such instances, courts must interpret and apply the [New York] Rules in a manner that ‘balances the varying federal interests at stake.’”

Looking at the specific facts of the case, the Second Circuit then found that applying New York’s confidentiality rule “could not have undermined the qui tam action” because the two other partners in FLPA, both nonlawyers, had sufficient information to bring the lawsuit. As a result, applying the New York rule to Bibi, the lawyer who had previously represented the company, “would not affect, much less undermine, the federal interests embodied in the FCA qui tam provision.” As for the New York rule itself, the Second Circuit found that this particular rule “implicitly accounts for the federal interests at stake in the FCA by permitting disclosure of information ‘necessary’ to prevent the ongoing commission of a crime.”

This “balancing of federal interests” aspect of the FLPA II opinion appears to contradict the court’s earlier finding that nothing in the FCA preempts state ethics rules. What would the Second Circuit have done if

---

310 FLPA II, 734 F.3d at 163. Although the relator’s brief to the Second Circuit did not use the term “preemption,” it attempted to avoid application of the New York Rules by citing federal authority to the effect that “[w]hen state rules of professional conduct governing attorneys would frustrate the purposes of federal law, a federal court should apply the Model Rules.” Brief and Special Appendix for Plaintiff-Appellant, supra note 207, at 37. The Model Rules permit disclosure of confidential client information not only to prevent future or ongoing crimes or frauds, but also to prevent, mitigate, or rectify the consequences of past crimes or frauds. See supra note 90 and accompanying text.

311 FLPA II, 734 F.3d at 163 (quoting Grievance Comm. v. Simels, 48 F.3d 640, 646 (2d Cir. 1995); United States ex rel. Dick v. Long Island Lighting Co., 912 F.2d 13, 18 (2d Cir. 1990)).

312 Id. (quoting Grievance Comm., 48 F.3d at 646).

313 Id. at 164–65 (“By FLPA’s own admission, it was unnecessary for Bibi to participate in this qui tam action at all, much less to broadly disclose Unilab’s confidential information. . . . ‘Baker and Michaelson each has ample relevant information to bring this case.’” (quoting Brief and Special Appendix for Plaintiff-Appellant, supra note 207, at 43)).

314 Id. at 165 & n.16 (“FLPA could have brought the qui tam action based on the information that Baker and Michaelson possessed as former executives of Unilab.”).

315 Id. at 164; see N.Y. RULES OF PROF’L CONDUCT r. 1.9(c) (N.Y. STATE BAR ASS’N 2015).
Bibi’s conduct had been governed by a confidentiality rule prohibiting *any* disclosure to prevent future or ongoing client crimes or frauds? In that situation, if the Second Circuit found that the state confidentiality rule did not adequately account for the federal interests at stake in the FCA, then the opinion suggests that the court would not have used such a rule to disqualify the relator, resulting in what looks very much like federal preemption of an otherwise applicable state rule. And with respect to the New York rules themselves, or similar rules, another court might find that a rule that permits disclosure only to *prevent* ongoing or future fraud unduly frustrates the purpose of the FCA, which provides financial incentives for disclosure that can *rectify* past fraud. And what about a lawyer’s loyalty obligations under state law? In *FLPA II*, the government took the position that if the Second Circuit agreed with the district court that the former-client conflict rule prevented the relator from filing a qui tam lawsuit, this decision “would drastically and negatively affect the impact of the *qui tam* provisions of the [FCA].”

Aside from the former-client conflict rule, state fiduciary law broadly prohibits both lawyers and nonlawyers from profiting from the disclosure of confidential client information, even when the disclosure itself is permitted—law that, if applied, would clearly have a devastating impact on the ability of insiders to serve as qui tam relators.

In *Holmes*, also an FCA lawsuit, the federal district court similarly rejected the relator’s argument that the FCA preempts state ethics standards. It quoted the *Doe* court’s observation that the FCA “does not authorize [a relator] to violate state laws,” and cited the *FLPA II* court’s rejection of federal preemption.

We believe that the possible preemption of at least some aspects of state ethics laws is a more complicated question than the *Doe, FLPA II*, or *Holmes* courts acknowledged. Although it is not our purpose to thoroughly explore this question, we must note some of the complicating factors, including, but not limited to, the “balancing of federal interests” test that the Second Circuit invoked in *FLPA II*. Other complicating factors include a

---

316 Brief and Special Appendix for Plaintiff-Appellant, *supra* note 207, at 15. The brief quotes the government as further stating: “The faults that Quest finds with Bibi—his close relationship to the company that allegedly engaged in the fraud and his disclosure of his inside knowledge of the fraud—are precisely what makes relators valuable sources of information.” *Id.*


318 *Id.* at *8* (quoting *Doe*, 862 F. Supp. at 1507).

319 *Id.* at *11* (citing *FLPA II*, 734 F.3d at 163). The *Holmes* court also noted that applying ethics standards to a qui tam relator like *Holmes* was particularly appropriate because *Holmes* was acting pro se—an option available to him only because of his status as a lawyer and not available to nonlawyer-relators.

320 *FLPA II*, 734 F.3d at 163.
number of FCA counterclaim cases involving nonlawyers, in which courts have held that FCA defendants may not assert certain state law claims, and a more complicated picture of how obstacle or frustration preemption has fared in the Supreme Court.

The FCA counterclaim cases involve attempts by company defendants to assert state law claims for indemnification and contribution as well as for enforcement of releases and nondisclosure agreements. For example, in 1991, in Mortgages, Inc. v. U.S. District Court, the U.S. Court of Appeals for the Ninth Circuit held that the district court should have dismissed seven state law-based counterclaims seeking indemnification and contribution. The Ninth Circuit concluded that the “FCA is in no way intended to ameliorate the liability of wrongdoers by providing defendants with a remedy against a qui tam plaintiff with ‘unclean hands;’” that there is no federal basis for permitting FCA defendants to bring counterclaims for indemnification or contribution; and that there is “no right to assert state law counterclaims that, if prevailed on, would end in the same result.” Similarly, in 2005, in United States ex rel. Longhi v. Lithium Power Technologies, the U.S. Court of Appeals for the Fifth Circuit ruled that a company could not seek to enforce release and indemnification clauses in a contract with the Department of Defense because enforcing such an agreement would “ignore the public policy objectives expressly spelled out by Congress in the FCA [and] provide disincentives to future relators.”

Courts considering FCA cases have also limited state law counterclaims that are closely analogous to lawyers’ confidentiality and loyalty duties, refusing to enforce fiduciary or contractual duties to inform the defendant internally of the alleged violation and limiting the availability of counterclaims based on contractual nondisclosure agreements.

322 Id. at 214 (emphasis added).
These counterclaim cases do not expressly refer to federal preemption of state law, but rather speak in the language of public policy. Nevertheless, they give preemptive effect to the FCA over conflicting state law obligations, using analysis that is analogous to obstacle or frustration preemption. And this form of implied preemption appears to be far more complex and controversial than the cursory treatment in *Doe* and *FLPA II* suggests.

According to one recent commentary, the presumption against preemption in cases involving the historic police powers of the states is merely a presumption—not a “clear statement rule” requiring the identification of language in the statute declaring the purpose of Congress to preempt state law. Indeed, the same commentary observes that in obstacle preemption cases, “courts have no text dealing with preemption to construe.” As a result, courts must evaluate “the degree of tension between state law and

---

2004) (“[P]ublic interest in the production of [defendant’s] records trumps any fiduciary obligation or any confidentiality agreement.”).

325 See, e.g., United States ex rel. Ruhe v. Masimo Corp., 929 F. Supp. 2d 1033 (C.D. Cal. 2012) (denying defendant’s motion to strike the relators’ exhibits that were allegedly taken in violation of a non-disclosure agreement, and noting that “the strong public policy [of the FCA] would be thwarted if [a defendant] could silence whistleblowers”).

326 At least two court decisions and one article have acknowledged the federal preemption frame for analyzing these counterclaim issues. *Siebert*, 2013 U.S. Dist. LEXIS 149145, at *13 (noting that “[i]t is possible . . . that the holding in *Mortgages* is based on conflict preemption principles”); United States v. Dynamics Research Corp., 441 F. Supp. 2d 259, 269 (D. Mass. 2006) (examining whether counterclaims are preempted by the FCA); Thomas F. O’Neil III et al., *The Buck Stops Here: Preemption of Third-Party Claims by the False Claims Act*, 12 J. CONTEMP. HEALTH L. & POL’Y 41, 51 n.53 (1995) (“Many of the cases [addressing contribution and indemnification] do not specifically discuss the preemption doctrine or explain their holdings in these terms. However, the courts’ rationale that state law claims would undermine the FCA amounts to the application of the doctrine sub silentio.”). The counterclaim cases could also be characterized as demonstrating preemptive effect of the federal common law. See Ernest A. Young, *Preemption and Federal Common Law*, 83 NOTRE DAME L. REV. 1639, 1640 (2008) (describing the development of “federal common law rules in cases concerning the rights and obligations of the United States”).

327 Ernest A. Young, “*The Ordinary Diet of the Law*: The Presumption Against Preemption in the Roberts Court,” 2011 SUP. CT. REV. 253, 271. Professor Young cites and quotes from Justice Scalia’s concurring and dissenting opinion in *Cipollone v. Liggett Group, Inc.*, a 1992 Supreme Court case also cited by the *Doe* court:

> See, for example, *Cipollone v. Liggett Group, Inc.*, 505 U.S. 504, 545 (1992) (Scalia, J, concurring in the judgment in part and dissenting in part) (“Though we generally ‘assume that the historic police powers of the States [are] not to be superseded by . . . Federal Act unless that [is] the clear and manifest purpose of Congress,’ we have traditionally not thought that to require express statutory text.”) (quoting *Rice*, 331 U.S. at 230).

*Id.* at 271 n.85 (alterations in original); see also *Doe*, 862 F. Supp. at 1507 n.12 (citing *Cipollone*, 505 U.S. at 545, as authority for federal preemption inquiries).

328 See Young, supra note 327, at 274.
congressional purpose,” and the critical question will be “just how much conflict is tolerable.” According to another commentary, “so-called ‘obstacle preemption’ potentially covers not only cases in which state and federal law contradict each other, but also all other cases in which courts think that the effects of state law will hinder accomplishment of the purposes behind federal law.” Finally, scholars have noted the Supreme Court’s failure to use the presumption in a consistent manner, prompting calls for its explicit abandonment and causing at least one commentator to declare that the Court has “created a presumption in favor of preemption.” Given the apparent weakening of the presumption against preemption, it is hardly surprising to find that FCA courts in nonlawyer cases have not hesitated to effectively preempt counterclaims based on state law in light of the obstacle such claims present to furthering the purposes of the FCA.

If the FCA counterclaim cases are correct that the FCA preempts some state law obligations based on contract or agency law, then does it also preempt a lawyer’s similar obligations, including those based on state attorney conduct rules? Not necessarily. Some commentators have argued that lawyers, together with other independent third parties, play a unique role as “gatekeepers,” whose “consent is necessary to allow issuers or investors to proceed with transactions,” or as “corporate monitors,” who monitor a company’s compliance with its legal obligations. Gatekeepers arguably do not need financial incentives because they already have strong “reputational concerns that stem beyond their relationship with the corporation.”

329 Id. at 274–75; see also Thomas W. Merrill, Preemption and Institutional Choice, 102 NW. U. L. Rev. 727, 729 (2008) (“[T]he key question in most preemption cases entails a discretionary judgment about the permissible degree of tension between federal and state law, a question that typically cannot be answered using the tools of statutory interpretation.”).

330 Caleb Nelson, Preemption, 86 Va. L. Rev. 225, 228–29 (2000). Professor Nelson is critical of such a broad use of obstacle preemption, concluding that “the mere fact that federal law serves certain purposes does not automatically mean that it contradicts everything that might get in the way of those purposes.” Id. at 231–32.

331 Young, supra note 327, at 307 & n.290 (quoting Mary J. Davis, Unmasking the Presumption in Favor of Preemption, 53 S.C. L. Rev. 967, 971 (2002)).

332 See, e.g., Pacella, supra note 7, at 1057–58.

333 Id. at 1058. This part of Pacella’s argument is more persuasive than her further argument that lawyers in general are “commonly described as ‘gatekeepers’” and therefore should also be prohibited from receiving financial incentives. Id. While some lawyers serve as gatekeepers, many do not. Pacella also describes lawyers as “corporate monitors,” with preexisting duties to the company to intervene to avoid corporate wrongdoing that could harm the company. Id. at 1057. As such, she argues that they “should not be incentivized in the same manner to disclose information as the rank-and-file employee,” presumably because financial incentives may cost them “their intimate relationship with the corporation.” Id. at 1058. Given that we do not believe that all corporate lawyers perform the narrow “gatekeeping” role, we suggest that the more general description of lawyers as “corporate monitors” could be used to buttress normative arguments that financial incentives to blow the whistle are incompatible with unique lawyer functions. See id. at 1057–58.
As for corporate monitors, their effectiveness depends on a company’s willingness to treat them as trustworthy insiders, which may be difficult once the company understands the lawyer’s eligibility to receive a whistleblower bounty.334

It may also be appropriate to treat lawyers’ obligations stemming from a state ethics code differently than obligations arising under general common law, including the law regulating private agreements. For example, as a matter of public policy, there is no reason to think that contractual confidentiality arrangements between two private parties serve the public interest, but a state supreme court’s adoption of a professional rule for lawyers does reflect that court’s assessment of the competing public interests at stake in a lawyer-client relationship.335 These include both the interest in protecting the confidentiality expectations of clients and the interest in preventing, mitigating, or rectifying client crimes or frauds.336 So there may be more reason to defer to a state professional rule than a contractual arrangement between private parties or even a fiduciary’s common-law obligation not to profit from a client’s confidential information.

In addition, if FCA courts were to routinely enforce nondisclosure agreements against relators who disclose fraud, that could spell the end of qui tam cases because companies can impose such agreements on all of their employees. Similarly, if FCA courts were to routinely enforce an employee-agent’s common-law duty not to profit from the employer’s confidential information, there would be few insiders in a position to serve as qui tam relators. Court enforcement of lawyer ethics rules, on the other hand, will affect only a narrow slice of possible relators. This excluded group would include lawyers currently representing clients in the same or a substantially related matter and former lawyers for a client where applicable state ethics codes prohibit disclosure under the particular circumstances. Depending on a court’s interpretation of the former-client conflict rule, this group could also include former lawyers bringing a qui tam suit against a former client in the same or a substantially related matter; however, under the Second Circuit’s “balancing of federal interests” approach in FLPA II, such a broad interpretation of an ambiguous ethics rule could fail on the

334 This normative argument against lawyer eligibility for whistleblower awards would also seem to apply to nonlawyer compliance professionals. In drafting its Dodd-Frank whistleblower regulations, the SEC specifically allows compliance professionals to be eligible for its whistleblower awards, but on terms different from other employees. They must first engage in internal whistleblowing and then wait 120 days before submitting information about violations to the SEC. 17 C.F.R. § 240.21F-4(b)(4)(ii)(B), (iv)(C) (2012).

335 See FLPA II, 734 F.3d at 154.

336 See id.
ground that it does not sufficiently take account of the federal interests under the FCA.337

B. Dodd-Frank

Like the FCA, the Dodd-Frank whistleblowing provision is silent on the question of federal preemption of state law, including state laws governing lawyers. Similarly, the SEC’s implementing regulations do not use the term “preemption.”338 But those regulations build on the agency’s earlier regulation of lawyers under SOX, which created new confidentiality exceptions for lawyers practicing before the SEC,339 and asserted that it preempts state ethics rules that are “in conflict” with the SOX regulation.340 When the SEC promulgated the SOX regulation, it argued that contrary state ethics rules would frustrate the purposes of SOX, whereas some elements of the organized bar argued that the SEC lacked the authority to preempt state confidentiality rules.341 Given that the SOX legislation expressly delegated to the SEC a mandate to set “minimum standards of conduct” for lawyers practicing before the SEC, we find the SEC’s arguments largely persuasive.342 Nevertheless, because the SOX legislation focused on up-the-ladder

337 See id. In refusing to give an ambiguous state court rule its broadest possible meaning, out of deference to important federal interests, it is unclear whether the district court in FLPA I was engaging in the functional equivalent of federal preemption. See FLPA I, 2011 U.S. Dist. LEXIS 37014, at *22–32.

338 On the other hand, the SEC’s intention to preempt conflicting state ethics rules in its whistleblower program is clear from the way it defines the “original information” that is eligible for a whistleblower award. 17 C.F.R. § 240.21F-4(b). Ordinarily, information obtained “in connection with the legal representation of a client” cannot form the basis for an award. Id. § 240.21F-4(b)(4)(ii). But the SEC permits such information to be the basis for an award if a lawyer may disclose such information pursuant to either the SOX regulation or “the applicable state attorney conduct rules,” id. § 240.21F-4(b)(4)(iii), and the SOX regulation itself asserts that it preempts conflicting state confidentiality rules, 17 C.F.R. § 205.1 (2003).

339 17 C.F.R. § 205.3(d)(2).

340 Id. § 205.1 (“Where the standards of a state or other United States jurisdiction where an attorney is admitted or practices conflict with this part, this part shall govern.”). The regulation also provides a safe harbor for lawyers facing state bar discipline for making disclosures consistent with the regulation. Id. § 205.6 (“An attorney who complies in good faith with the provisions of this part shall not be subject to discipline or otherwise liable under inconsistent standards imposed by any state or other United States jurisdiction where the attorney is admitted or practices.”).

341 See Roger C. Cramton et al., Legal and Ethical Duties of Lawyers After Sarbanes-Oxley, 49 VILL. L. REV. 725 (2004), for an insightful summary of the debate between the SEC, the ABA, and the Washington and California state bars.

342 15 U.S.C. § 7245 (2012). The statute directs the SEC to issue “minimum professional standards, including” the specific standards on reporting up within organizations. Id. The use of the word “including” indicates that Congress’s grant of authority to the SEC is broader than simply the reporting up rules, and the use of the word “minimum” indicates that the SEC regulations
reporting, and did not expressly contemplate lawyers reporting violations outside of the company to the SEC, we acknowledge that the preemptive status of the SEC’s SOX regulation is at least questionable.343

The issue of SOX preemption was never judicially resolved, perhaps because no lawyers made disclosures that were permitted by that regulation but prohibited under applicable state rules.344 But now that the SEC has promised financial incentives for individuals—including lawyers—who make whistleblower disclosures under Dodd-Frank, lawyers may (finally) be disclosing information that they have been permitted to disclose since 2003. Even so, it is not clear how a lawyer’s disclosure under the Dodd-Frank program will result in a judicial determination of whether SOX preempts more restrictive state confidentiality rules, given that the statute requires the SEC to keep whistleblowers’ identities confidential.345

The SEC’s Dodd-Frank regulations clearly state that lawyers are not precluded from receiving an award if their disclosure was permissible either under state attorney ethics codes or under the SOX regulation.346 But even if lawyers are ethically permitted to disclose client wrongdoing to the SEC, what about their loyalty obligations under state law? What is the significance of the SEC’s failure to even mention lawyers’ loyalty obligations under state law?347 A lawyer-whistleblower might argue that the SEC regulation, read literally, permits a lawyer to receive an award, so long as the disclosure itself is permitted.348 But it is unclear whether, in drafting the regu-
lation, the SEC ever considered conflicting ethical obligations other than those found in confidentiality rules, and it is similarly unclear whether it intends to grant an award if it is convinced that the lawyer is violating a state law that is not preempted by the prior SOX regulation.

Here, as with the FCA, if courts clearly confront the preemption question, they will need to seriously consider the possibility that some aspects of the state regulation of lawyers unduly frustrate the objectives of the Dodd-Frank whistleblower program. Once again, some commentators believe that some lawyers might be distinguished from nonlawyers on the basis of their role as “gatekeepers” or “corporate monitors.” We believe that lawyer conduct rules can be distinguished from contract and general fiduciary law that applies to lawyers and nonlawyers alike. If so, then the SEC should refuse to reward a lawyer who clandestinely discloses client information in an ongoing representation, given that the lawyer would be violating code provisions prohibiting current client conflicts and requiring keeping the client informed of significant developments. On the other hand, the SEC might conclude that broader state fiduciary law is preempted, because preventing lawyers and other fiduciaries from profiting as a result of their disclosure of client information could eviscerate the whistleblower reward program that Congress has authorized.

IV. CHOICE OF LAW

As we have seen, state ethics rules frequently differ, particularly rules governing the disclosure exceptions to the duty of confidentiality. If nei-

that provision. See id. Lawyers who acquire information as a result of a client representation are mentioned only in the provision defining “[o]riginal information,” and, as previously described, that section clearly treats as “original” any information that a lawyer is ethically permitted to disclose under either state attorney conduct rules or the SOX regulation. See supra note 338 and accompanying text.

349 The ABA submitted a letter to the SEC arguing that financial incentives for whistleblowing lawyers “create[] an objectionable conflict of interest.” Pacella, supra note 7, at 1049 n.135 (quoting Letter from Stephen N. Zack, President, Am. Bar Ass’n, to Sec. & Exch. Comm’n (May 20, 2011), http://www.americanbar.org/publications/governmental_affairs_periodicals/washington letter/2011/june/attorneyclientprivilege.html [http://perma.cc/Q65C-29N7]). However, the SEC commentary accompanying both the preliminary and final Dodd-Frank regulation does not mention lawyers’ loyalty obligations, including conflicts of interest and keeping the client informed of important developments.

350 See id. at 1057–58.

351 We refer to current and not former lawyers because unlike the FCA, a Dodd-Frank lawyer-whistleblower is not acting on behalf of another; therefore, a former lawyer is not even arguably violating the former-client conflict rule. Similarly, unlike current lawyers, a former lawyer generally has no obligation to keep the client informed of developments in the subject of the representation.

352 See supra notes 26–123 and accompanying text.
ther the FCA nor Dodd-Frank preempts state ethics rules, then lawyers contemplating blowing the whistle need to know which state’s rules apply. Unfortunately, given that these whistleblower programs often involve national companies with multiple offices, as well as in-house lawyers who are not necessarily licensed in either the state where they advise the company or the state where a disclosure may occur, many lawyers will have difficulty predicting which state’s ethics rules govern. As with preemption, we do not intend to thoroughly explore the choice of law issues raised in these federal programs. Rather, our more limited purpose is to note the complexity of the problem and the failure of the existing case law to grapple with this complexity. In addition, because the SEC’s Dodd-Frank regulation apparently preempts state confidentiality rules that are more strict than the SEC’s own SOX regulation, and because loyalty provisions do not differ significantly from state to state, we will focus our discussion on the very difficult choice of law issues that arise when a federal court attempts to determine the ethical propriety of a lawyer-relator’s disclosure of confidential client information in bringing a qui tam lawsuit under the FCA.

In the United States ex rel. Doe v. X Corp. series of cases, the in-house lawyer who sought to serve as a qui tam relator was licensed in Pennsylvania, but worked for the defendant company in California and in Virginia, where he was terminated. When he informed the company that he was going to file a retaliatory discharge lawsuit, the company filed an action in federal district court in Virginia, seeking the return of documents and a preliminary injunction preventing him from disclosing the company’s confidential information.

Doe argued that because he was a member of the Pennsylvania Bar, Pennsylvania ethics code provisions should apply. The court, however,
found that “a [federal] court sitting in a diversity case applies the conflicts [of law] rules of the forum state” and that “Virginia applies the substantive law of the place of the wrong,” citing a Virginia torts case. Because Doe lived in Virginia at the time of the lawsuit, had taken copies of documents and maintained them there, and because “the confidential disclosures sought to be restrained would presumably be made in or from Virginia,” the court applied the Virginia Code of Professional Responsibility and issued the injunction that subsequently prevented Doe from receiving an award from the settlement of the FCA lawsuit he had filed. The district court’s findings as to choice of law were upheld without discussion in Doe’s appeal to the U.S. Court of Appeals for the Fourth Circuit.

Given that Doe was a diversity action and not an FCA lawsuit, the district court understandably relied on conflict of laws rules applied in diversity actions, thereby looking to the forum state’s conflict of laws rules. The court did not explain, however, why it chose to look at the conflict of laws rules applied in torts cases, as opposed to disciplinary cases, or the conflict of laws rules applied in contracts, agency, or other lawsuits involving the provision of services. If the Virginia Code of Professional Responsibility had contained a choice of law provision, then perhaps the court would have looked to that provision instead. At that time, however, neither the Virginia Code nor other state attorney codes had a special choice of law provi-

---

358 X Corp. I, 805 F. Supp. at 1304 n.11 (citing McMillan v. McMillan, 253 S.E.2d 662 (Va. 1979)).
359 Id.
360 Id. at 1312; see Under Seal v. Under Seal, 17 F.3d 1435, 1435 (4th Cir. 1994) (per curiam) (affirming the district court’s grant of injunctive relief).
361 Under Seal, 17 F.3d at 1435 (“While John Doe was qualified as a lawyer in Pennsylvania, at the time he was terminated by X Corp. and for some time previously, he was an X Corp. employee in Virginia.”). In a subsequent opinion granting summary judgment for X Corp. on its various claims and on Doe’s claim for retaliatory discharge, the district court cited its earlier opinion for the proposition that “Virginia law governs here” and that “Doe advances no new or sufficient reason to reconsider this conclusion.” X Corp. v. Doe (X Corp. II), 816 F. Supp. 1086, 1091 n.8 (E.D. Va. 1993), aff’d sub nom. Under Seal, 17 F.3d 1435.
363 X Corp. I, 805 F. Supp. at 1304 n.11. In stating that “Virginia applies the substantive law of the place of the wrong,” the court cited—without explanation—a case that involved a torts claim. See id. (citing McMillan, 253 S.E.2d 662).
sion. It was only in 1993 that the ABA Model Rules were amended to provide such a provision, and that provision was amended again in 2002. As a result, different state ethics rules now have not only different confidentiality and other substantive provisions, but also different approaches to choice of law for purposes of attorney discipline.

In United States ex rel. Fair Laboratory Practices Associates v. Quest Diagnostics, Inc. (“FLPA”), the relator filed a qui tam action under the FCA. The company filed a motion to dismiss the lawsuit and to disqualify the partnership-relator—Fair Laboratory Practices Associates (“FLPA”)—as well as its general partners and its counsel on the ground that Bibi, one of the general partners, had violated his ethical obligations as a former in-house lawyer for the company by disclosing confidential information in connection with the lawsuit. Defendant Quest was a publicly traded Delaware corporation headquartered in New Jersey. Bibi was licensed only in New York, but had been employed in New Jersey by defendant Unilab, before Unilab was acquired by Quest, and was still working in New Jersey at the time of the lawsuit. The two other former Unilab executives who joined Bibi to form FLPA were also employed at that time in New Jersey. Shortly after

---

365 See Geoffrey J. Ritts, Professional Responsibility and the Conflict of Laws, 18 J. LEGAL PROF. 17, 20 (1993) (noting that the 1993 amended version of Model Rule 8.5 was the first codified choice of law rule for lawyer conduct).

366 Id.

367 See, e.g., 2 HAZARD & HODES, supra note 152, § 70.05 (discussing the revision of Rule 8.5(b) as part of the ABA’s Commission on Multijurisdictional Practice project).

368 AM. BAR ASS’N, STATE IMPLEMENTATION OF ABA MJP POLICIES (Aug. 24, 2014), http://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/recommendations.authcheckdam.pdf [http://perma.cc/6P3C-PJ6A] (demonstrating that thirty-five states have adopted the current version of Model Rule 8.5(b) or substantially similar provisions, two states and the District of Columbia have rules based on the 1993 version of Model Rule 8.5(b), and thirteen states have a different rule or no rule at all); CPR POLICY IMPLEMENTATION COMM., AM. BAR ASS’N, VARIATIONS OF THE ABA MODEL RULES OF PROFESSIONAL CONDUCT: RULE 8.5: DISCIPLINARY AUTHORITY; CHOICE OF LAW (May 4, 2015), http://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/mrpc_8_5.authcheckdam.pdf [http://perma.cc/LIG7-2BUG].


370 Id. at *17–19.

371 Id. at *2–4.

372 Id. at *7.

373 See Brief and Special Appendix for Plaintiff-Appellant, supra note 207, at 8, 38 n.30.

374 FLPA I, 2011 U.S. Dist. LEXIS 37014, at *42. At the time they formed FLPA, Bibi and the other former executives were employed by yet another company, Life Sciences Research, Inc., located in Hackensack, New Jersey. Id. at *5.

375 Id.
the partnership was formed, FLPA filed the original qui tam lawsuit in a federal district court in New York.  

The federal district court in FLPA I took a different approach than the Doe court to the choice of law issue, although it ultimately reached the same result of applying the ethics rules of the forum state. It began by noting that “[w]hile federal courts may look to the Model Rules of Professional Conduct and state disciplinary rules for guidance, such rules are not binding on this court.” It went on, however, to note that “[t]he ‘salutary provisions [of the New York rules] . . . have consistently been relied upon by courts of this district and circuit in evaluating the ethical conduct of attorneys.” The court then looked to the choice of law provision in the New York Lawyer’s Code of Professional Responsibility, which stated:

For conduct in connection with a proceeding in a court before which a lawyer has been admitted to practice (either generally or for purposes of that proceeding), the rules to be applied shall be the rules of the jurisdiction in which the court sits, unless the rules of the court provide otherwise.

Because Bibi was a member of the New York Bar and the FCA lawsuit had been filed in the Southern District of New York, the court concluded that the choice of law provision in the New York Code dictated that Bibi’s conduct be evaluated under the relevant provisions of the New York Code. As in Doe, the circuit court of appeals upheld the district court’s choice of law determination without significant discussion.

FLPA I is arguably more instructive than Doe because it involved a lawsuit brought under the FCA. As a result, the court was not bound to ap-

---

376 Id. at *17.
377 Id. at *19 (quoting Hull v. Celanese Corp., 513 F.2d 568, 571 n.12 (2d Cir. 1975)).
378 Id.
381 FLPA II, 734 F.3d at 163 (quoting Hull, 513 F.2d at 571 n.12) (noting merely, without acknowledging any dispute over choice of which state’s laws to apply, that “[a]s a general matter, the ‘salutary provisions [of New York’s ethical rules] have consistently been relied upon by the courts of this district and circuit in evaluating the ethical conduct of attorneys’”); see also X Corp. I, 805 F. Supp. at 1304 n.11. The Second Circuit FLPA II opinion noted that the district court had applied the New York Lawyer’s Code of Professional Responsibility, which had been replaced by the New York Rules of Professional Conduct after the filing of the lawsuit. FLPA II, 734 F.3d at 157 n.1. Because the parties relied on the New York Rules in their briefs, and because the parties agreed that the substantive standards were the same, the Second Circuit’s opinion cites the New York Rules throughout its opinion. Id. The court did not reference the fact that FLPA had argued in its brief that the New Jersey Rules, not the New York Rules, should apply to Bibi’s conduct. See Brief and Special Appendix for Plaintiff-Appellant, supra note 207, at 38 n.30.
ply the forum state’s choice of law rules, as it would be in a diversity action, because state law did not directly apply to the federal question of FLPA’s eligibility to serve as a qui tam relator.\footnote{Neither the Doe nor the FLPA I court was bound to apply the rules of professional conduct of any jurisdiction. These rules are adopted for the purpose of providing a basis for lawyer discipline and are not binding on courts in other proceedings, including claims involving malpractice, breach of fiduciary duty, disqualification, and fee forfeiture. See, e.g., \textit{Model Rules of Prof’l Conduct} scope 20. In other types of proceedings addressing a lawyer’s conduct, courts frequently look to rules of professional conduct, but violation of an attorney conduct rule is not necessarily determinative outside the context of a disciplinary proceeding. See \textit{id}. Of course, in Doe, a diversity action, the court was bound to apply state law, including state precedent as to the role of disciplinary rules in breach of fiduciary duty actions. See supra note 304 and accompanying text. In FLPA I and II, however, we argue that the courts were not bound to look to state law at all, but rather could rely solely on federal law to determine the eligibility of relators for a financial award under the FCA. See infra note 401 and accompanying text. Nevertheless, having chosen to rely on state law, which the courts were free to do, the FLPA courts needed to determine which state’s rules to apply. \textit{Cf. X Corp. I}, 805 F. Supp. at 1304 n.11 (applying choice of law principles in determining which state’s law applied to the company’s former lawyer in diversity case).} Although it was not bound to do so, the court looked to the law of the forum state, in this case New York, because New York’s professional conduct rules “have been consistently relied upon by courts of this district and circuit in evaluating the ethical conduct of attorneys.”\footnote{\textit{FLPA II}, 734 F.3d at 154.} Here the court cited both Second Circuit precedent involving disqualification of a lawyer representing one of the parties, and the federal district court’s own local rule, which was adopted from the New York Lawyer’s Code of Professional Responsibility.\footnote{\textit{FLPA I}, 2011 U.S. Dist. LEXIS 37014, at *19–20 (citing Southern District of New York Local Civil Rule 1.5(b)(5)). The Second Circuit appellate opinion similarly applied the New York ethics rules, but without any discussion of the choice of law issue. See \textit{FLPA II}, 734 F.3d at 163. The district court’s local rule governed the discipline of attorneys for conduct “[i]n connection with activities in this Court,” Southern District of New York Local Civil Rule 1.5(b)(5), but Bibi was not appearing before the district court as an attorney representing a client; therefore the applicability of the local rule was questionable. See \textit{FLPA I}, 2011 U.S. Dist. LEXIS 37014, at *19–20. Similarly, we argue that the district court’s reliance on Second Circuit precedents in attorney disqualification cases was also questionable. See \textit{id}.} It is true that in ruling on a motion to disqualify a lawyer, many courts more or less automatically choose to apply the rules of the forum or, in the case of federal district courts, the rules of the forum state.\footnote{Like the Southern District of New York, most federal district courts have adopted local rules for lawyer conduct, and like the Southern District, most local rules adopt the lawyer conduct code of the forum state. See Judith A. McMorrow, \textit{The (F)Utility of Rules: Regulating Attorney Conduct in Federal Court Practice}, 58 SMU L. REV. 3, 10–12 (2005).} But this approach is problematic. The application of the forum’s own rules to attorneys representing the parties in a lawsuit is probably justified by courts’ inherent authority to regulate the conduct of the lawyers who appear before them, as well as the distinction between procedure and substance, under which “the
forum will apply its own local law to matters of procedure and the otherwise applicable law to matters of substance.” 386 But in FCA lawsuits, a lawyer-relator is not appearing before the court as an attorney representing a party, 387 and the ethical propriety of the relator’s conduct does not involve matters of judicial administration or related procedural questions for which the “forum has compelling reasons for applying its own rules.” 388 Rather, we believe that the ethical conduct of a lawyer-whistleblower under the FCA is best characterized as substantive because it affects the lawyer’s eligibility to receive a whistleblower award. 389

386. Restatement (Second) of Conflict of Laws, ch. 6, intro. note (Am. Law Inst. 1971); cf. Comm. on Prof’l Responsibility, Ass’n of the Bar of the City of N.Y., Uniform Ethics Rules in Federal Court: Jurisdictional Issues in Professional Regulation, 50 Rec. Ass’n B. City N.Y. 842, 870 (1995) (stating that rules for “[l]awyers litigating in federal court should . . . be . . . in harmony with important objectives of the judicial system”). According to the Restatement (Second) of Conflict of Laws, the distinction is between issues involving “judicial administration,” in which “[t]he forum has compelling reasons for applying its own rules to decide such issues even if the case has foreign contacts and even if many issues in the case will be decided by reference to the local law of another state.” Restatement (Second) of Conflict of Laws § 122 cmt. a.

387. In 2015 in United States ex rel. Holmes v. Northrop Grumman, the federal district court similarly relied on federal precedent in addressing motions to disqualify a lawyer from representing a client in federal litigation. See No. 1:13cv85-HSO-RHW, 2015 U.S. Dist. LEXIS 71804, at *9 (S.D. Miss. June 3, 2015) (citing Fifth Circuit choice of law precedent in disqualification cases). In that case, however, Holmes was appearing pro se, and the court gave as its reason for relying on such precedent that lawyer-relators (and not nonlawyer-relators) are entitled to appear pro se because they have ethical obligations, including obligations of candor to the court, and violations of these obligations “may result in serious consequences to the errant attorney.” Id. at *5 (quoting United States ex rel. Schwartz v. TRW, Inc., 118 F. Supp. 2d 991, 995 (C.D. Cal. 2000)). We believe that most lawyer-relators will choose to be represented by counsel; therefore, federal district courts are likely to rely primarily on FLPA I and II in addressing choice of law issues. Unlike the Second Circuit and most other circuit courts of appeal, the Fifth Circuit looks not to the ethics rules of either the district court or the forum state, but rather to both “applicable ‘local and national ethical’ standards.” Id. at *9 (quoting FDIC v. U.S. Fire Ins. Co., 50 F.3d 1304, 1314 (5th Cir. 1995)). In identifying applicable “local” standards, the Mississippi federal district court looked not only to the Mississippi rules, but also to the D.C. Rules because much of Holmes’s conduct occurred while the case was pending in D.C. federal district court. Id.

388. Restatement (Second) of Conflict of Laws § 122 cmt. a.

389. There are also some other settings in which a court could properly apply its local rules to a lawyer-party who seeks to represent others, but not in a lawyer-client relationship, such as when a lawyer-plaintiff seeks to serve as a class representative or to pursue a shareholder derivative action. In those cases, the court must determine whether the lawyer-plaintiff will serve as an adequate representative, therefore, the issue is properly characterized as procedural. See, e.g., Fed. R. Civ. P. 23(a)(4) (class action); id. r. 23.1(a) (shareholder derivative action). Certainly in federal courts, it is well-settled that ethical issues that arise in the context of applying the Federal Rules of Civil Procedure are best settled under federal procedural law because of the significance of upholding federal law and policy. See, e.g., Boccardo v. Comm'r, 56 F.3d 1016, 1019 (9th Cir. 1995) (holding that the state ethical rule forbidding lawyers from assuming litigation costs conflicted with Rule 23 of the Federal Rules of Civil Procedure and was therefore invalid); Rand v. Monsanto Co., 926 F.2d 596, 600 (7th Cir. 1991) (finding that the local rule limiting advancing of legal costs is trumped by Federal Rules of Civil Procedure because local rules adopting state ethical
Of course, the FLPA I court looked initially to the New York Code not for its substantive provisions, but rather for its choice of law rule. But this too was problematic. As with most disciplinary choice of law rules, the New York Code had two provisions: one for litigation and one for non-litigation. The New York litigation choice of law provision looks to the rules of the jurisdiction in which the court sits for lawyer conduct “in connection with a proceeding in a court before which a lawyer has been admitted to practice” (either generally or for purposes of that proceeding) . . . unless the rules of the court provide otherwise. But there was no indication that Bibi had been admitted to practice before the federal district court where the lawsuit was pending.

Unlike the New York rule cited in FLPA I, the Model Rules’ litigation choice of law provision applies the rules of the jurisdiction in which a tribunal sits “for conduct in connection with a matter pending before a tribunal,” without regard to whether the lawyer has been admitted to practice before that tribunal. Nevertheless, we believe that “conduct in connection with a matter pending before a tribunal” should generally be confined to the conduct of a lawyer representing a party, and not to the conduct of a lawyer-relator. This is because when lawyers are appearing before the tribunal as party-representatives and “officers of the court” the forum has the greatest interest in applying its own rules.

---

391 See N.Y. LAWYER’S CODE OF PROF’L RESPONSIBILITY DR 1-105(B)(1) (litigation choice of law provision); id. r. 1-105(B)(2) (non-litigation choice of law provision).
392 Id. r. 1-105(B)(1) (emphasis added); see also N.Y. RULES OF PROF’L CONDUCT r. 8.5(b)(1) (N.Y. STATE BAR ASS’N 2015) (same). For a discussion of the non-litigation provision of the New York Code and Rules, see infra notes 397–399 and accompanying text.
393 See MODEL RULES OF PROF’L CONDUCT r. 8.5(b)(1). For a discussion of the non-litigation provision of the Model Rules, see infra note 400 and accompanying text.
394 See MODEL RULES OF PROF’L CONDUCT r. 8.5(b)(1).
395 See, e.g., id. r. 3.3 cmt. 2 (providing that the rule concerning “Candor Toward the Tribunal” establishes “the special duties” of a lawyer to “act[] as an advocate in an adjudicative proceeding” and to “avoid conduct that undermines the integrity of the adjudicative process”).
396 There is another problem with applying the Model Rules’ litigation choice of law provision. Consider a situation in which a lawyer-relator’s only connection to the state in which the court sits is that it is where the FCA lawsuit was filed. If the Model Rules’ litigation choice of law rule applies, then the lawyer can easily avoid restrictive state confidentiality rules by filing the FCA lawsuit in a state with favorable confidentiality exceptions, even when that state is not where the lawyer is licensed, where the lawyer principally practices, or where any of the pre-litigation conduct occurred. Under the FCA, a relator may bring an action “in any judicial district in which the defendant . . . can be found, resides, transacts business, or in which any act proscribed by [the
In *FLPA I* itself, the court’s apparent mistake in applying New York’s litigation choice of law provision was immaterial because application of the non-litigation choice of law provision would have achieved the same result: that provision dictates that “[i]f the lawyer is licensed to practice only in this state, the rules to be applied shall be the rules of this state,” and Bibi was licensed only in New York. But what if Bibi was not licensed in New York? Such a situation is simply not contemplated under the New York choice of law rule, which assumes that the lawyer is licensed at least in New York and makes no provision for disciplining a non-New York lawyer, particularly one whose conduct has its predominant effect in a jurisdiction other than one in which he or she is licensed. This approach, which is followed in several other jurisdictions, makes sense in lawyer discipline cases, where the adopting jurisdiction is free to choose not to attempt to discipline non-admitted lawyers. But it makes little sense in other contexts, including federal court determinations of the eligibility of a lawyer-relator as a result of possibly unethical conduct.

The Model Rules’ non-litigation choice of law provision does not suffer from this particular problem: indeed, it makes no reference to where the lawyer is licensed to practice. Rather, Model Rule 8.5 provides that, for non-litigation conduct, the rules to be applied are “the rules of the jurisdiction in which the lawyer’s conduct occurred, or, if the predominant effect of the conduct is in a different jurisdiction, the rules of that jurisdiction shall be applied to the conduct.” If this rule had been adopted in New York and applied in *FLPA I*, then the court could have determined that the relevant conduct occurred either in New Jersey (where Bibi had been employed by Unilab and where he had presumably first disclosed confidences to his FLPA partners) or in New York (where Bibi disclosed confidential information in his sealed

397 N.Y. LAWYER’S CODE OF PROF’L RESPONSIBILITY DR 1-105(B)(2); see N.Y. RULES OF PROF’L CONDUCT r. 8.5(b)(2) (same); see also *FLPA I*, 2011 U.S. Dist. LEXIS 37014, at *19–20 (discussing choice of law provision of DR 1-105(B)(1)).

398 See N.Y. RULES OF PROF’L CONDUCT r. 8.5(a)–(b) (discussing disciplinary authority for lawyers who are licensed to practice within the state).

399 New York follows the approach of the 1993 version of Model Rule 8.5, which is also followed by the District of Columbia. See id.; see also D.C. RULES OF PROF’L CONDUCT r. 8.5 (D.C. BAR ASS’N 2015); MODEL RULES OF PROF’L CONDUCT r. 8.5(b) (AM. BAR ASS’N 1993) (failing to include a provision regarding choice of law for a lawyer who is not licensed to practice within the state).

400 MODEL RULES OF PROF’L CONDUCT r. 8.5(b)(1) (“A lawyer shall not be subject to discipline if the lawyer’s conduct conforms to the rules of a jurisdiction in which the lawyer reasonably believes the predominant effect of the lawyer’s conduct will occur.”). This is another aspect of the disciplinary rules that makes sense in the disciplinary context but not necessarily in other contexts.
complaint to both the court and the government, and where further disclosures in connection with the lawsuit might be anticipated). To the extent that the “predominant effect” of the conduct might have occurred elsewhere, the court might have looked to the state where the defendants were located, which would have been either New Jersey, where Quest was headquartered, or California, where Unilab had been headquartered before becoming a wholly owned subsidiary of Quest.

If a federal court hearing FCA lawsuits is not compelled to look to the choice of law rules of the forum state to determine the ethical propriety of a qui tam relator’s conduct—and we believe it is not—then we suggest that the court should then be free to apply whatever choice of law rule it deems is best suited to the issue at hand, as an application of federal common law.401 Rather than start from scratch, we also suggest that the court adopt as federal law either the Model Rule choice of law provision for non-litigation matters402 or the “most significant relationship” test adopted by the Restatement (Second) of Conflict of Laws.403 Under the Restatement’s provision for identifying the state with the most significant relationship with respect to the rights and duties between agents and their principals,404 the

---

401 In disqualification cases, federal courts have long had a tendency to create common law and develop responses “largely unconstrained by formal rules.” McMorrow, supra note 385, at 3; see also Eli J. Richardson, Demystifying the Federal Law of Attorney Ethics, 29 GA. L. REV. 137, 156–57 (2005) (explaining that even when local district court rules provide for an exclusive source of law for attorney ethics, courts often rely on other standards). For a general discussion of the role of federal common law in filling gaps in federal statutes, see Henry J. Friendly, In Praise of Erie—And of the New Federal Common Law, 39 N.Y.U. L. REV. 383 (1964).

402 For a discussion of why the Model Rules’ non-litigation provision is superior to the litigation provision on the question of the eligibility of an FCA relator to receive an award, see supra notes 397–401 and accompanying text.

403 RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 6. Under the Restatement (Second) of Conflict of Laws, courts are instructed to balance a number of factors in choosing “the applicable rule of law.” Id. This balancing test is generally described as an attempt to determine the state with “the most significant relationship” to the issue of law in question. See, e.g., RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 5 cmt. h (AM. LAW INST. 2000) (providing that the Restatement (Second) of Conflict of Laws’ general balancing test seeks to determine “the rule with the most significant relationship to the charged offensive conduct,” and is similar to the “predominant effect” principle ushered in by revisions to Model Rule 8.5(b) in 1993); see also RESTATEMENT (SECOND) OF CONFLICT OF LAWS, § 6 cmt. c (stating that the goal of balancing test is to “state a general principle, such as application of the local law ‘of the state of most significant relationship,’ which provides some clue to the correct approach but does not furnish precise answers”).

404 Balance-of-factor tests are notoriously unpredictable; however, the Restatement (Second) of Conflict of Laws gives more specific guidance in provisions addressing different subject areas in which a dispute may arise, including tort, contract, or agency law. See supra note 403 and accompanying text. In our view, agency law best captures the fiduciary nature of the relationship between lawyer and client, and it is precisely the fiduciary nature of the lawyer-client relationship that underlies the rules on confidentiality and loyalty. See, e.g., Ritts, supra note 365, at 62–63 (in
court would then look to the substantive ethics rules of the state where the lawyer performed as an agent, or, if the performance occurred in several states, to the state where the lawyer most frequently acted on the client’s behalf.405 The two tests are similar,406 but the Restatement approach may be preferable as it avoids the need to determine whether the “predominant effect” of the lawyer’s conduct will be in a jurisdiction other than where the conduct occurred—a determination that is fraught with indeterminacy.

CONCLUSION

Whether lawyers should be permitted to seek financial rewards for blowing the whistle on their clients is a difficult normative question that we have not addressed in this Article. Answering that question will require exploration of the nature of the relationship between the lawyer and the client and consideration of whether lawyers occupy a role that is significantly different than other company insiders, who also owe obligations of confidentiality and loyalty to the company. It will also require having a clear understanding of the extent of the obligations that lawyers owe to their clients under current legal doctrine, including any special obligations to entity clients. This is the descriptive project that we have undertaken.

In this Article, we have explored the four doctrinal questions that are key to determining whether a lawyer may receive a whistleblower award under a federal program without violating state fiduciary law and state rules of professional conduct: (1) Under what circumstances may a lawyer disclose a client’s confidential information to others? (2) Under what circumstances does a lawyer’s obligation of loyalty preclude acting adversely to a client, including seeking personal benefit when engaging in conduct that is permissible for other purposes, such as to prevent or rectify harm to another? (3) Are any of a lawyer’s obligations under state law preempted by federal law that provides for financial incentives for whistleblowers? (4) Which state’s law applies to lawyers who move from state to state as they work for national companies?

applying the Restatement (Second) of Conflict of Laws to a matter involving successive representation and a former government lawyer, the “issue could have been cast as one of agency law, with the ethics rules regulating an attorney/agent’s fiduciary duties of loyalty and confidentiality”). But see White Consol. Indus., Inc. v. Island Kitchens, Inc., 884 F. Supp. 176 (E.D. Pa. 1995) (applying the general provision of Restatement (Second) of Conflict of Laws section 6 to a lawyer’s professional responsibility question); David A. Barry & William L. Boesch, Massachusetts Legal Malpractice Cases, 93 MASS. L. REV. 321, 339–40 (2011) (arguing for the application of the Restatement (Second) of Conflict of Laws’ provisions on contracts for the rendering of services).

405 RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 291 & cmt. f.
406 See RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 5 cmt. h (referring to the 1993 version of Model Rule 8.5(b)).
As the reader has already discerned, these issues turn out to be enormously complex. For example, state confidentiality exceptions differ widely among the states, and it is not clear whether or when disclosures in the course of seeking a financial reward are “reasonably necessary” either to prevent the perpetuation of continuing crimes or frauds or to rectify or mitigate past wrongdoing. State loyalty obligations include not only conflicts of interest rules for current and former clients—rules that may be difficult to apply in the context of FCA and Dodd-Frank whistleblowers—but also a lawyer’s common-law fiduciary duty not to profit from a client’s confidential information. This fiduciary duty applies to both current and former clients and is so broad that it would appear to prohibit any lawyer from seeking a financial reward for disclosing client information without the client’s consent.

Given the potential breadth of a lawyer’s confidentiality and loyalty obligations, we question the view of the three court decisions that concluded without much discussion that the FCA does not preempt a lawyer’s ethical obligations under state law. None of these courts confronted the fact that federal courts in FCA cases involving nonlawyers have apparently preempted the state law fiduciary or contractual duties of nonlawyers if enforcing their state law obligations would effectively prevent them from serving as FCA relators. If these are indeed preemption cases, as we believe they are, then federal courts will need to distinguish lawyers from nonlawyers if the similar obligations of lawyers are not also to be preempted in at least some situations. Finally, there is no clear choice of law rule that federal courts must apply in determining which state’s ethics laws apply to a lawyer-relator when there are several states with significant contacts with the matter. We recommend that federal courts develop their own federal common-law choice of law rule for FCA cases, but we also acknowledge the difficulty of formulating a choice of law rule that is both principled and predictable in its application.

Unfortunately, only a few courts have addressed the issues we discuss in the context of the FCA, and none have done so in the context of Dodd-Frank. Moreover, as a practical matter, it is unlikely that there will be many additional court decisions addressing whether and when lawyers can obtain financial incentives under these statutes. We anticipate few additional FCA suits brought by lawyer-relators against their former clients because the apparent confidentiality and loyalty constraints discussed above make such suits less likely to succeed than those filed by nonlawyers. Because relators’ lawyers are generally paid on a contingent basis, a rational relators’ lawyer will be less likely to accept and pursue a lawyer-relator’s FCA case than one brought by a relator not limited by these ethical constraints. The key gatekeeper for potential lawyer-relators is the relators’ bar, and they are likely to
recognize the greater risk they would be taking by agreeing to represent a lawyer-relator.

The Dodd-Frank whistleblower program is also unlikely to produce court decisions on these issues. This is because the key decision-maker under this statute is the SEC rather than the judiciary. The SEC will be (and perhaps already is) making its decisions about whether to grant whistleblowing awards to lawyers through a secret proceeding that lacks an adversary presentation of issues and is largely insulated from judicial scrutiny. Thus, a company that might view itself as harmed by a lawyer-whistleblower’s disclosure to the SEC may never have an opportunity to contest the lawyer-whistleblower’s initial tip or eventual award.407

Despite our reluctance to address the normative questions involved in lawyer whistleblowing for financial reward, the ability of a Dodd-Frank whistleblower to maintain anonymity throughout the process leads us to conclude by raising yet another normative question. Even if financial rewards for lawyer-whistleblowers are appropriate in some circumstances, is a secret process, shielded from judicial review, ever appropriate for determining whether a particular lawyer is eligible to receive a whistleblower reward?

It may be appropriate to grant lawyers financial incentives for blowing the whistle on a former client. But Dodd-Frank’s process for granting these awards—without an adversary presentation of the issues and without judicial oversight—is a cause for concern. Another government program, asset forfeiture, arguably has similar parameters: a secretive administrative process through which local law enforcement agencies seize property without any adversarial presentation of issues or evidence, with the burden on property-owners to seek judicial review of such seizures.408 The record of the asset forfeiture program demonstrates both the power of using financial in-

---

407 As a formal matter, the SEC’s award decisions are subject to judicial review only if the SEC denies an award entirely or grants an award below the 10% statutory minimum. 17 C.F.R. § 240.21F-13(a). As a practical matter, the respondent company may never know that the SEC’s investigation was triggered by a whistleblower’s tip or the identity of that whistleblower. Even though the SEC issues press releases regarding its whistleblower awards, those releases provide little specific information about the whistleblower, the respondent company, or the securities violations involved. See 15 U.S.C. § 78u-6(h)(2)(A); In the Matter of the Claim for Award in Connection with [Undisclosed Party], supra note 260. Therefore the administrative process for granting whistleblower awards is unlikely to generate a matter subject to judicial review.

centives to enforce legal norms and the danger of combining those incentives with a secret administrative process.409