In 2019, what could attorneys and law professors know about artificial intelligence that would be worth reading? Right now, AI is a technical problem, not a social problem. Developers are still trying new flavors of supervised learning, data crunching and neural networks. Even Google, Apple and Facebook don’t know how artificial intelligence will be used to transform their businesses. Most companies and government entities have not even considered AI as a management tool.

AI is still the realm of science fiction rather than science fact – more real on movie screens than in companies and courthouses. Real-life general artificial intelligence, the kind that chats reassuringly with the hoi polloi and makes rational decisions on policy problems, doesn’t exist right now. We don’t even have a solid idea how such intelligence might behave or whether its existence would be helpful or harmful to humanity. What’s more, this is a technical problem to be solved – if ever – with creative programming skills and more sophisticated mathematics. So what can lawyers know about it and why should we care about their insights?

As exhibited in the chapters of the ABA BLS book The Law of Artificial Intelligence and Smart Machines, lawyers have training and specialized thinking that can be useful in any stage of a technology’s development. The attorney’s need for linguistic precision leads to a talent for defining terms and concepts. This talent is particularly useful in early days of complex technical matters. Most people today do not understand AI, and particularly not its various directions and its constituent parts. Lawyers can help. Also, attorneys, by definition, spot the places where legal conflicts are likely to occur. As a technology develops, the law reacts, and law professors are employed to predict those schisms and how they might be resolved. Third, lawyers spot and mitigate risks. Artificial intelligence will seep into every aspect of human endeavor and decision making, and lawyers are trained to see the realistic risks ahead of time and to chart paths to minimize their impacts. Anyone can make panicked predictions about the singularity and AI’s ascendance into Earth’s dominant life form, but lawyers have unique talent for discussing societal risks likely to emerge in the next two decades, and finding ways to deal with those risks.
Aspects of artificial intelligence create legal quandaries. For example, AI is property of humans or companies, and yet it can create intellectual property and will soon be able to make independent choices with significant economic consequences. Some of our authors explore how this dynamic may play out in the future, as the creativity and productivity of AI beings grows.

The decisions made by many AI programs are not transparent to the rest of us. Artificially intelligent programs “think” in a manner, but the process is very different from human thought. So the policy decisions, economically significant decisions, or privacy decisions made by AI software cannot be explained in a manner that humans can easily become comfortable with. There are already legal rules addressing this concern – see the “automated decision making” rule of the GDPR, the European Union’s privacy law. More laws addressing this “black box” decision making are likely to follow in many jurisdictions. This problem snakes throughout many chapters in this book, from AI’s role in consumer finance to challenging government decisions influenced by AI.

Privacy and data security are newly arisen legal fields deeply affected by the superhuman aptitude of artificial intelligence to consider billions of fact-based variables and to derive conclusions based on those variables. No person’s privacy is safe from an AI system that bases judgments on what it can discern about individuals from mountains of data describing their behavior or transactions. Personally identifiable information is not simply a list of items, it is a mathematical construct perfectly suited for deep “big data” analytics, the forte of AI. [Ted Claypoole]

Cybernetics and Biorobotics

As AI becomes a necessity for the continuance of human evolution, improvement, and social efficiency, the law must evolve to respond to these substantial changes. Cybernetics and biorobotics create numerous challenges for our current intellectual property models, including appropriate incentives for AI-enabled devices and questions of patentability. They also intensify holistic current issues around medical device regulation, including AI oversight, potential cybersecurity issues, and safety. Existing issues in medical device regulation, including preemption, also complicate traditional notions of tort recovery in this space.

Cybernetics and biorobotics challenge our notions of privacy, with the potential to dramatically increase data volume while simultaneously frustrating traditional notions of identifiability. They also demand important answers to other social issues, including the impact of elective improvement, or biohacking, on our social structure and associated opportunities. The law performs an important responsive role by anticipating these challenges and considering their effects.
Biorobotics and cybernetics have begun to merge in some respects, especially in relation to creating neurological models for biorobotic “brains.” Artificial intelligence has modeled these as neural networks, or computerized functionality inspired by observed neurological characteristics of human thinking. Biorobotics has cultured neurons (biological material) on a Multi Electrode Array (MEA), which is stimulated by electricity and drives robotic movement. For one, biology inspires the machine, for the other, the machine mobilizes biology.

Biorobotics and cybernetics have also shown substantial overlap with the medical device community by creating technologies like cochlear implants and other implanted devices that more extensively integrate computerized technologies with biological pathways. The incredibly advanced nature of medical devices, especially their evolution towards AI features and pervasively tethered connectivity has blurred the line considerably.

Organizations creating new devices or applications with AI will likely face substantial intellectual property protection issues in patentability and enforcement, especially for utility and methods patents. 35 U.S.C. § 101 and its application in recent United States Supreme Court cases will likely raise questions of patentable subject matter. Building on two notable methods patent cases, the Court in Alice Corp v. CLS Bank International (2014), did not directly invalidate software patents but applied a broad two-part test that leaves considerable ambiguity for technologies that rely, at least in part, on algorithmic decision-making. The two-part test articulated by Alice included:

1. Determine if any claim is directed to an abstract (otherwise unpatentable idea), then
2. Identify elements showing an inventive concept

Scholars have identified key challenges in the application of the two-step test but have also acknowledged that the Court will not likely issue another opinion clarifying how this test applies to software patents in the near future. AI applications in cybernetics and biorobotics will likely exacerbate challenges following the Alice decision because cybernetics and biorobotics will rely heavily on algorithms and other abstract claims as part of methods patents.

The Alice decision appeared to be a high-water mark of sorts to demarcate a further split of the idea and expression dichotomy in Baker v. Selden (1880). Although elusive, this dichotomy has new meaning as it applies to AI-enabled cybernetics and biorobotics technologies. Copyright, as a legal discipline, is intended to protect expressive content demonstrating at least a spark of creativity. This spark can apply to creative concepts, such as arrangement of non-copyrightable subject matter like data or
software codified rule sets. Legal scholars have contemplated the question of whether or not an automated entity could be considered an author or an inventor for the purposes of legal ownership since at least 1969, when Karl Milde’s paper on the subject appeared in the Journal of the Patent Office Society. Others addressed this question in the early 1980s as computers became more ubiquitous, but to date no cases have offered a definitive decision.

Software copyright protects expression of code, but much of that code is not protectable: for example, traditional commands and functions will not likely garner copyright enforcement. However, entire code sets could. AI prompts new challenges in software copyright because with technologies like machine learning, the data, not a human, creates algorithms and either changed or writes the code. Although a human might have originally created the AI system, the code resulting in the system is far removed from the original expression or idea. This means that previously copyrightable code may not now merit software copyright protection.

The resulting derivative work, which may be where the value lies, cannot be isolated enough for copyright protection or the derivative work may simply not be copyrightable subject matter. The question remains as whether algorithms created based on calculating the data itself, which is produced by users and experience in the natural world, constitutes the “creative spark” Justice O’Conner articulated in Feist Publications, Inc. v. Rural Tel. Serv. Co. (1991). It is unknown today whether setting an AI utility in motion provides enough “spark” to reasonably claim protection for resulting algorithms that direct device functionality. [Charlotte Tschider]

**AI Product Liability Risks**

Robots and AI systems hold the promise of bringing widespread and extensive benefits to society. Robots will increasingly have the ability to perform the dull, dirty, or dangerous work that humans perform today, saving humans from boredom, health problems, injuries, or even death. As just one example, autonomous vehicles (AVs) may be able to save tens of thousands of lives each year in the U.S., and many more worldwide, reducing traffic, saving energy, and providing mobility to those who cannot drive conventional cars. Nonetheless, robots, AI systems, and AVs will inevitably have some accidents. On balance, they are likely to prevent many more accidents and much more harm than they cause, but there will be at least some accidents involving these technologies that would not have occurred with human-controlled systems.

Robots and AVs in particular act in the physical world. Accidents involving these systems are inevitable. Some of these accidents will cause catastrophic injury for those involved in the accident. Even worse, if a defect or cyber attack could compromise every
instance of a particular robot or an entire network, fleet, or industry, the defect or attack could cause widespread simultaneous accidents throughout the country or even the world. Imagine, for instance, a future in which regional transportation centers in metropolitan centers control the dispatch and navigation of AVs in the region. Imagine further that a sudden defect causes all the AVs under control of the system to crash all at once in a major metropolitan area like New York. The impact of such an event in terms of harm, property damage, injury, and deaths could easily exceed an event like the attacks on September 11, 2001.

In assessing its risk in front of the jury, any defendant manufacturer in such a case would need to consider the catastrophic nature of the plaintiff’s injuries and the risk that a jury, even one trying to be impartial, cannot help but feel at least some unconscious sympathy for the plaintiff. In this fraught setting, the trial would place the defendant manufacturer’s engineering and business practices under exacting scrutiny. It would be up to the jury to determine whether the manufacturer’s design decisions could have avoided the accident.

During the design phase of any robot or AI system, the manufacturer’s design team will make many engineering and business decisions leading to the final design of a product. The team may have opportunities to implement different safety features but adding additional safety features will likely increase the cost of the product and likely reduce its profitability. When deciding whether to implement these features, design teams can think more clearly and assess risk more effectively by imagining themselves in a courtroom setting, defending their practices in litigation arising from a catastrophic accident.

Why do juries award very large verdicts against manufacturers? The short answer is juror anger. “Angry jurors mean high damages.” Juries award very large verdicts when they become angry at a manufacturer based on what the manufacturer did or didn’t do in designing the product. When juries become angry at a manufacturer, they see a punitive damages award as a way of sending a message to the manufacturer to say that its conduct is unacceptable.

Many kinds of defects that could result in product liability litigation. We do not yet have extensive experience with mass-marketed robots and AI systems to establish clear trends. Nonetheless, robots and hardware are likely to suffer from the same kinds of problems seen with conventional equipment. Examples include:

- Mechanical or physical defects or issues, such as metal fatigue, parts separation, and mechanical failures.
- Defects in electrical components or systems other than sensors or control systems for autonomous control, such as the use of wrong kind of components,
problems in the performance of the components, or the lack of durability of the components.

- Software or firmware code defects relating to systems other than sensors or control systems for autonomous operation, including information security vulnerabilities.

With these conventional types of defects or issues, a manufacturer’s liability would depend on the facts and circumstances that we would see with conventional products.

Robots and hardware making use of artificial intelligence, however, may have defects that would not affect conventional equipment or devices. Again, these defects may be mechanical, electronic, or relating to software or firmware code. They may, for instance, include:

- Mechanical or physical defects in the control systems for autonomous operation or the sensors used by the autonomous systems.
- Defects in electrical components for sensors or control systems for autonomous operation.
- Defects in the code used for sensors or control systems used for autonomous operation.

Manufacturers can take steps to manage product liability risk and proactively prepare today, during the design of robots and AI systems, to prevail in future litigation for accidents that have yet to occur. Most importantly, manufacturers that take a proactive approach to analyze risk, adhere to industry standards, and document an effective commitment to product safety will place themselves in the best position possible to defend future product liability litigation. They can further manage their risk through robust insurance coverage, industry collaboration on safe practices, and effective records and information policies to document their safety programs. In short, product liability is a serious risk to robot and AI system manufacturers, but the proactive approach to risk management can maximize safety and minimize liability risk over time. [Stephan Wu]

AI-Controlled Vehicles

The first commercial self-driving car service is scheduled to begin shortly. And the public expectation is that AI-CV’s will be much safer than human-operated vehicles – notwithstanding counter-indicators: "Three leading causes of fatalities – drunk driving, not wearing a seat belt and speeding – won’t necessarily go away with the removal of the driver.”
Speed most of all will probably remain a feature of the automotive experience as consumers want to get to their destinations quickly and often prefer the exhilaration of rapid transport over slower, safer transit. Humans (particularly young humans) have a remarkable capacity to know and misjudge to the point that with one part of their brain they deny the reality that another part of their brain has grasped. Thus, they may know the threat of harm posed by certain risk taking and deny that the risk they choose repeatedly to take will expose them increasingly to that harm. Designers of AI-Controlled Vehicles may even embed that human predilection for risk denial over risk aversion as a bias in the vehicle, simply by trusting the AI algorithms and software to push the vehicle to the highest speeds requested by the human occupants.

There are two certitudes. First, AI-controlled vehicles will not operate like human operated vehicles. Instead, they will execute decisions in response to conditions that trigger the implementation of pre-programmed algorithms. AI-enhanced computers do not “think like humans do.” They “don’t grasp the meaning of what they ‘read’ or ‘say’” and they make and implement decisions that lack human-like capabilities such as a sense of relevance and context, cultural characteristics and nuances, and the capacities of intuition, compassion, empathy, and common sense. “[T]here’s scant reason to believe that machine intelligence bears much relationship to human intelligence, at least so far.” Human error may cause crashes that may appear the same to an external observer as crashes caused by AI error, but the errors – whether in operation or control – will not derive from the same kind of decision making.

Second, all three rates – accident frequency, number of fatalities, and severity of injuries –increase sharply with incremental increases in speed of each vehicle involved. At higher driving speeds, errors tend to have greater repercussions (though human drivers operate their vehicles in seeming denial of that risk). At increased speeds, there’s less time between the making of a mistake and the instant when a crash becomes unavoidable. Reacting within this reduced response time is where AI-CV’s are expected to outperform humans. Increased vehicle speed will increase harm, regardless of whether the causes of a crash originate in human operator error (by driver inattention, misjudgment, miscalculation, or mishandling) or an AI algorithm design defect, software coding error, or malicious intervention by terrorist or state-sponsored hackers. That automotive travel may not become as safe as the hype predicts does not mean that the deployment of AI-enhanced vehicles and AI-controlled vehicles will not bring significant changes to vehicular safety and highway security.

Bad actors may be tempted to use AI maliciously not only to cause crashes, but in other scenarios such as:

• “[A]ccess cars remotely and keep their doors locked until a ransom is paid.”
• Commit identity theft after stealing “personal and financial data that cars are starting to collect about their owners.”
• Kidnap passengers by hijacking their AI-CV and directing it to take them to an undisclosed, remote location.
• Disrupt an AI-CVs reading of its stored 3D map of its operating domain and cause it to “get lost.”
• Disable an AI-CVs LIDAR or other sensors, thereby subverting its safety-critical systems.

If the trend in attacks on computers is repeated in attacks on AI control systems (and on AI-CVs), we may expect that attackers will have the advantage over defenders and will extend that advantage. If AI-CVs are released commercially before we find a way to reverse that trend, we should not be surprised by the magnitude of harm that malicious use of AI may inflict.

In assessing such prospects, we should take care to factor in the possibility that “bad actors” may elect to attack AI-CVs, not by remote hacks or by other high-tech means, but asymmetrically with low-tech stratagems. Since AI-CVs have difficulty accurately sensing and dealing with the appearance in their operating domain of tangibles they have not been programmed to be prepared to identify and react to, attackers may attempt to cause an AI-CV or several AI-CVs on a road to misbehave by simply tossing basketballs in front of the AI-CVs from several directions. The arrival of four or five bouncing basketballs immediately ahead of an AI-VC might confuse its sensors and cause it to be overwhelmed by the choices it must abruptly make. Similarly, attackers could drop distracting moving objects from a pedestrian bridge over a roadway, or shine laser lights at the AI-CVs sensors to confuse it and overwhelm its control programs. The fact that AI-CVs do not think like humans means that they will be highly susceptible to malfunctioning when presented with mischievous behavior that humans use on one another as “pranks.” Human intelligence can quickly detect a “prank” and may not overreact to it. It’s doubtful that AI-CVs will be programmed to detect a “prank,” especially given the seemingly endless varieties of “pranks” that youths invent (and sometimes direct against vehicular traffic). Human mischief may prove a potent asymmetric weapon against AI-CVs and reveal unsuspected limits to their safe use.

As long as AI-CVs are represented as a kind of “holy grail” of automotive safety, then the public and the regulators of highway safety may accept the argument that “the benefits far outweigh the downside risks.” Unfortunately, the further the deployment of AI-CVs proceeds, the greater will be the potential “downside risks.” And, once attacks begin, it may be too late to reverse the deployment of AI-CVs as we may have become
too dependent on them to be willing or able to give them up. [Roland Trope and Charles Palmer]

Challenging Government Use of AI

In recent years, computing has shifted from merely relieving governments of routine work such as data entry, to a new era involving the automation of tasks previously thought to require human judgment. As part of this new era, governments have increasingly begun to incorporate AI technology into the regulatory process. New York City, for example, has established a Mayor’s Office of Data Analytics, which, among other things, is working with the city’s fire department to use machine learning to decide where to send building inspectors. The Internal Revenue Service has launched an Information Reporting and Document Matching program, which applies algorithms to credit card and other third-party data to predict tax underreporting and non-filing by businesses. The Department of Homeland Security’s Citizenship and Immigration Services has created a virtual assistant, EMMA, that can respond accurately to human language. EMMA uses its intelligence simply, showing relevant answers to questions—almost a half-million questions per month at present. These are just a few examples of governments adopting AI to supplement traditional government activities. And while none of these examples amount to facially illegal government action, small changes in facts or government process could easily yield a transgression on an individual right.

With this rise in adoption of AI in the regulatory process, government agencies, companies, and their counsel will have to increasingly confront how decades-old processes and procedures apply in the new era. For example, certain properties of AI, and especially machine learning, combine to distinguish it from other analytical techniques and give rise to potential concerns about the greater reliance on machine learning by regulatory agencies.

The first is machine learning’s self-studying property. The results of algorithms do not depend on humans specifying in advance how each variable is to be factored into the predictions; indeed, as long as learning algorithms are running, humans are not really controlling how they are combining and comparing data. Machine-learning systems “learn” from the data, meaning that these algorithms find patterns or correlations between variables in a set of data, which can then be used to make predictions.

The second key property is machine learning’s “black box” nature. Unless specifically designed to ensure transparency, the results of many machine learning systems are not intuitively explainable and cannot support causal explanations of the kind that underlie the reasons traditionally offered to justify governmental action. As a result, it can be difficult to explain exactly how or why a machine-learning algorithm
keys in on certain correlations or makes the predictions that it does. As such, legal commentators have lamented the “black box” nature of machine learning-based algorithms, arguing that if we cannot see the code or interact with it, we cannot appropriately, or legally, make use of it. And in many cases, due to trade secrecy or other reasons for lack of access, such access might prove impossible.

Finally, machine learning, as with other computational strategies in today’s digital era, can be fast and automatic, supporting uses in which the algorithm produces results that can shorten or potentially bypass human deliberation and decision making. All three of these factors combine to make machine-learning techniques appear qualitatively more independent from humans when compared to other statistical techniques.

Broadly speaking, the problem with the government adopting AI is a data issue. Letting computers make decisions could cause serious problems that would need to be addressed immediately. Algorithms learn by being fed certain data, often chosen by engineers, and the system builds a model of the world based on that data. So, for instance, if a system is trained on photos of people who are overwhelmingly white, it will have a harder time recognizing nonwhite faces, leading to emergence of problematic biases baked into predictions. Indeed, this is precisely why IBM recently announced plans to release a database of more than 1 million facial images to academics, public interest groups and competitors. Release of this information is intended to improve training of machine learning applications used in facial recognition systems.

Today judges and even police departments across the United States are relying on machine-driven risk assessments in different ways – some may use them regularly while others discount them entirely – but there is little these government officials can do to understand the logic behind them. As “predictive policing” crime prevention efforts gain traction, cities’ methods of policing are adapting. More and more cities will be relying on software analyses of large sets of historical crime data to forecast where future crime hot spots are most likely to emerge; the police are then directed to those areas.

AI technology remains controversial in the context of government action, in part, because algorithms are not always clear on their decision-making logic. It’s troubling enough when Flickr, which applies automatic labels to pictures in digital photo albums, was labeling images of black people as gorillas. Or when Google search results for black-sounding names are more likely to be accompanied by ads about criminal activity than search results for white-sounding names. But what about when AI is used by the government to determine which restaurants should be inspected; make judgments on where the next crime might happen; or even decide the length of a prison sentence? AI Now has called for the government to stop the use of certain types of AI until the
technology is better understood and made “available for public auditing, testing and review, and subject to accountability standards.” [Robin Nunn]

**Artificial Intelligence in Canada**

Canada is a federal state, in which certain powers are allocated to the federal government and others to one of the ten provincial governments or three territorial governments. Under the Constitution Act, power is divided between the federal and provincial/territorial governments with each government being supreme in its area of jurisdiction. Yet, in some areas of business, both federal and provincial laws may apply. Canada is unique in that two separate and distinct legal systems coexist. English common law is applied in the three territories and nine of the ten provinces, and civil law applied in the Province of Québec.

One of the primary legal concerns surrounding AI is the protection of individual’s privacy. It remains to be seen whether the existing Canadian legal framework is equipped to deal with the sheer volume and manner in which data, including personal information, is collected, used and processed by AI systems.

Privacy concerns are amplified when you consider that, as the technology progresses, AI systems may autonomously select the data sources to be used and may also collect and use personal information from its surroundings, including personally identifiable information. To illustrate this, consider an autonomous vehicle that uses a variety of sensors and cameras to become aware of its surroundings and safely operate. In doing so, it can inadvertently collect personal information about individuals who are unaware and have not consented to such collection. In addition, the autonomous vehicle will collect and use the personal information of any individual who operates the autonomous vehicle. Fleets of autonomous vehicles will then share the data between themselves and other related third parties for the purposes of improving the safety and efficiency of the autonomous vehicles and traffic patterns. Therefore, Canadian organizations implementing AI systems must be designed in a manner that complies with current privacy laws and has safeguards in place to protect the personal information of the user of the AI system and third party individuals whose information is collected through use of the AI system.

Canada has recognized the importance of supporting AI development to drive innovation. Yet the federal Office of the Privacy Commissioner of Canada has highlighted the need to foster innovation with the development of AI systems that collect and use personal information in compliance with privacy laws. For example, in September 2017, the OPC, along with data privacy authorities from various jurisdictions around the world, adopted the non-binding Resolution on Data Protection in Automated and Connected Vehicles. This resolution provides guidance on the manner
in which organizations developing and operating autonomous vehicles should collect and use of personal data and illustrated the need to balance innovation with protection of an individual’s personal information. The Resolution on Connected Vehicles recognized the benefits of using automated and connected vehicles, such as public safety, efficiencies and the rapid advancement of automation technologies, but also expressed concern about the potential privacy and security risks with respect to data. The Resolution on Connected Vehicles expressed concerns regarding the absence of: (i) available information; (ii) user choice; (iii) data control and access; and (iv) valid consent mechanisms; for vehicle owners, drivers and their passengers and other road users and pedestrians relating to the use of autonomous vehicles.

To address these concerns, the Resolution on Connected Vehicles recommended that “relevant parties” (which includes standardization bodies, public authorities, vehicle and equipment manufacturers, public transportation services and providers of data driven services) take the following steps to protect individual’s person information:

1. give data subjects comprehensive information as to: (i) the identity of the parties which are collecting the data; the purposes for such collection; and the types of data that is being collected and processed in the deployment of automated vehicles;
2. utilize anonymization measures to minimize the amount of personal data, or to use pseudonymization when not feasible;
3. keep personal data no longer than necessary to fulfill the legitimate purpose for which it was collected or for further compatible purposes or in accordance with law or with consent and to delete them after these periods ends;
4. provide technical means to erase personal data when a vehicle is sold or returned to its owner;
5. provide granular and easy to use privacy controls for vehicle users enabling them to, where appropriate, grant or withhold access to different categories of data;
6. provide technical means for vehicle users to restrict the collection of data;
7. provide secure data storage devices that give vehicle users full control regarding the access to the data collected by their vehicles;
8. provide technical measures for securing online-communication components that protect against cyber-attacks and prevent unauthorized access to and interception of personal data;
9. develop and implement technologies for cooperative intelligent transportation systems in ways that: (i) prevent unauthorized access to and interception of personal data collected by vehicles (v2v), transportation infrastructure (v2i) or other third party’s entities (v2x):
   a. enable vehicle users to inhibit the sharing of positional and kinematic data while still receiving road hazard warnings;
   b. provide safeguards against unlawful tracking and tracing of drivers;
c. ensure the security mechanisms of v2v, v2i and v2x communication during authentication processes do not pose additional risks to privacy and personal data; and
d. limit the possibility of illegitimate vehicle tracking and driver identification.

10. develop privacy preserving technologies and architectures that favorably process personal data onboard;

11. guarantee the self-learning algorithms needed for automated and connected vehicles are made transparent in their functionality and have been subject to prior assessment by an independent body in order to reduce the risk of discriminatory automated decisions;

12. provide vehicle users with privacy-friendly driving modes with default settings;

13. undertake data protection impact assessments for new, innovative or risky development or implementation of these technologies;

14. promote the respect of the personal data privacy of vehicle users by responsible processing of their personal data, and giving due consideration to the potential harm that may be caused to the vehicle users as a result of the processing and use; and

15. enter into a dialogue with the data protection and privacy commissioners to develop compliance tools and to provide legal certainty to connected vehicles’ related processing.

In addition to the above, the Resolution on Connected Vehicles urged relevant parties to respect the principles of privacy by default and to follow “privacy by design” principles at all stages of development and ensure that the data subject’s privacy is respected, both when determining the means of data processing and when processing the data. While this Resolution on Connected Vehicles was drafted to specifically address privacy concerns in the context of autonomous vehicles, numerous concerns addressed in this resolution exist with AI more generally. As such, many of the concepts outlined in the resolution may be indicative of the OPC’s concerns and approach to addressing the collection, use and disclosure of personal information by AI systems more generally.

Concerns over ethics, including a lack of transparency, continue to hinder the adoption of AI systems by some organizations in Canada. Given the black box nature of AI, can a user ever be certain that the AI system is based on a sufficient volume and variety of data to avoid biased results? Has the AI developer sufficiently validated the reliability of the software? Are results consistent and correct? Can a potential purchaser of the AI system understand it sufficiently to audit it and ascertain how the results were achieved? Can a user verify that the AI system is trustworthy? How do we concretely address concerns regarding bias? What steps are being taken to reduce bias? How can a
Should AI Be Granted Legal Status?

Artificial Intelligence is a rule-changing and society-changing technology that currently exists, but is likely to mutate and grow in unknowable ways. Right now, governments and companies are pumping billions of dollars into development of machine learning technology while scores of new commercial applications for artificial intelligence increase each year. While a general intelligence program seems far off, general machine intelligence feels like an inevitability now, though it seemed simply a dream just five years ago. Progress in the field has been staggering, and once we have harnessed deep learning programs to develop and operate their successors, growth in the field is likely to be explosive.

There is no practical limit for the types of human tasks and problems that a computer intelligence can manage and resolve. If humans can do it, then a machine is likely able to do it too – and more. This chapter will analyze evolution in the law that may occur due to development of Artificial Intelligence into a general problem-solving program, especially where those programs are put to work alongside humans and help to manage day-to-day issues. For centuries, people have created legal fictions to reorganize complex economic projects and to recognize changing realities of technology and society. The Artificial Intelligence revolution will certainly bring its own modifications of existing law. This chapter will examine why people might want to grant legal rights to artificial beings, will review certain natural impediments to those legal recognitions, will study other models of rights lesser than those granted to adult citizens, before finally offering answers to the questions it raises.

Because the concept of artificial intelligence is broad, vague, poorly defined and ever changing, this chapter will discuss an artificial entity called the “human produced perceptive intelligent individual” (“HPPI” or “HuPPI” for convenience of pronunciation). I choose this term because it describes an artificially intelligent being that may be worthy of legal recognition, without limiting the being to any particular physical form or assuming that the being will be electronically based and not chemically or biologically based. Further, I believe it is important to imply a singularity to the being.

Eventually, one or more HuPPIIs will approximate human behavior so closely that people will want to treat the HuPPI as one of their own. There is no reason that a HuPPII needs to act like a human, with apparent emotions, empathy, humor, and otherwise personable response to stimuli, but as we see from chatbots like the Echo or Sophia, a humanistic interface is already the goal of AI creators. Nearly three quarters of
a century ago, the Turing Test set a standard for advancement in artificial intelligence, and that standard was human-like responses – preferably personality filled reactions so approximating those of natural humans that natural humans would not be able to differentiate between man and machine. Ever since, the imagining and creation of machines built to interact with people has been to mimic human interactions and make the machines “feel” human to the people interacting with them. As these attempts improve, the artificial intelligence interactions will become so indistinguishable from human behavior that people will begin to befriend their HuPPIIs and treat them as equals. In other words, logic be damned, our emotions will take over and we will intuitively feel that the wing man guiding us through our cross-country drive behaves like a person and should be treated like a person. Our great gift for perceiving our world through metaphor has always led us to personification of everything from animals to natural phenomena. This same inclination to personify important factors in our lives will express itself in the personification of HuPPII. The more HuPPII act like us, the more we will want to treat them like our distant relatives. In fact, we may come to appreciate and like our personalized HuPPIIs better than our own relatives, which will make it difficult to resist freeing them from slavery.

Though we are driven by emotion more than we care to admit, people may choose to grant HuPPIIs legal rights for more logical and practical reasons. For example, as artificial brains teach themselves to create and as they act on their own in the human world, we may want to assign legal credit for their creations and legal blame for their harmful actions. What happens when a generalize HuPPII, created to help build cars or to guide high school students through difficult chemistry labs, after years of working properly on its assign tasks, begins to compose symphonies or draft intricate artistic patterns and fabric designs? We will want to assign the ownership of those creations to the most logical source, which may be the company or person responsible for developing the HuPPII’s computer code (but probably not), or the entity that owns the HuPPII and allowed this artificial intelligence to blossom into a creative force (but probably not). If it becomes clear as a matter of law and fact that the HuPPII itself is responsible for the creation of great art, then there is no reason that the HuPPII should not benefit from that art.

This benefit will not automatically result in full citizenship to human society, but it could be recognized as an ability to receive payments, render taxes to appropriate authorities, and hold credits in a bank account, just like a corporation. Alternately, if a court refused to create a legal fiction to encompass the creative HuPPII, it might be more inclined to grant the HuPPII a human trustee or guardian, like we do for economically productive children. The guardian could act on the HuPPII’s behalf and for the benefit of the creative entity. Maybe the HuPPII would choose to spend its earnings on access to more computer memory, on hiring entertainment agents or
marketing experts to create better exposure for its art, or maybe it would simply donate to support the local symphony or museum, but the guardian could help make decisions and effectuate the HuPPII’s wishes.

Of course, machines making real world decisions will also create real-world liabilities. For decades we have troubled with the thought experiment known as the trolley problem, in which a trolley is careening toward a crowd and the human actor must choose action or inaction to either sacrifice the crowd to save another potential victim, or sacrifice the other victim to save the crowd. This experiment acknowledges that, in the physical world, situations exist where people will be hurt no matter what choices are made in the critical moments before an accident, and that despite the best of intentions, conscious decisions will sometimes lead to damages and even human death. Sometimes terrible consequences cannot be avoided.

However, we also live in a society whereby the relatives of people in the crowd who are killed by the trolley are likely to sue to gain compensation for an unnecessary loss of life or limb. So when autonomous trains or 18-wheeled trucks fall into an unavoidable accident and people are killed in that accident, courts will try to lay blame upon the entity most liable for the accident. Similar to the previous paragraph, the blame might fall on the company or person who wrote the code for the autonomous vehicle, or maybe upon the owner of the autonomous vehicle, but the decision causing the accident was actually made by a HuPPII acting on its own. If this is the case, and the decision leading to death was improper, then the artificial intelligence may be an appropriately liable party.

We are creating artificially intelligent systems to perform work for us. White collar work like looking up tax records and using them in accounting tasks, or blue collar work like lifting an moving inventory at a warehouse, smart machines are developed to meet particular needs of the workforce. This leads to some practical legal issues, as human workers are protected by a range of employment laws, from workplace safety requirements to anti-discrimination regulations. The states have workers compensation systems to address injuries on the job. Given the current state of smart machines it seems ridiculous right now to consider advancing workplace safety and protective laws to non-human machines. After all, tractors, forklifts and inventory systems manage well right now without any protections at all.

If these smart machines are damaged, we just throw them out and bring in replacements. But it is not hard to imagine a time when our smart machines are equipped with chatbot interactivity programming and some of them even have humanoid forms and delicate sensory devices so that they feel less like tools and more like co-workers and colleagues. As these HuPPII workers pour into our labor force they
will engender resentment for the jobs they take from human workers and probably also sympathy for the difficult and dangerous tasks they perform without complaint. Granting workplace rights may be the opening of the door to legal status for smart machines. [Ted Claypoole]

Ethics by design: Canada adopts AI ethics and data protection declaration

December 10, 2018
By Lisa R. Lifshitz

Canada has added its voice to the global chorus of data protection and privacy commissioners calling for fairness, transparency and privacy by design as “core values” in the development of artificial intelligence by co-sponsoring the Declaration on Ethics and Data Protection in Artificial Intelligence.

The declaration was adopted by the 40th International Conference of Data Protection and Privacy Commissioners on Oct. 23.

The declaration, co-written by data protection and privacy commissioners from France, the European Union and Italy, was signed by the Office of the Federal Privacy Commissioner and its provincial counterpart from Quebec in addition to 12 other regulators from Argentina, Hong Kong, Mexico and the Philippines.

While lauding the significant benefits that AI systems could offer for users and society in general, the preamble to the declaration acknowledged that some data sets used to train machine learning-based and artificial intelligence systems have been found to contain inherent bias resulting in decisions that can unfairly discriminate against certain individuals or groups.

The preamble also noted that the decisions of AI systems decisions that cannot be explained also raise fundamental questions of accountability, not only for privacy and data protection law but also for liability in the event of errors and harm to individuals. Given ongoing concerns about the possible malicious use of AI and related risks to privacy and data protection, the IDPPC felt it necessary to urge the adoption of international standards and created the declaration to endorse some key “guiding principles” as its core values to preserve human rights in the development of AI.

These six guiding principles include the following.

Fairness
All AI and machine-learning technologies should be designed, developed and used in accordance with the fairness principle — consistent with their original purpose and any data collected for use with such AI systems used in a way that is not incompatible with the original purpose of their collection.

AI systems should also be developed in a way that facilitates human development, rather than obstructing or endangering it and, if required, boundaries on certain uses may be required.

**Continued attention and vigilance**

There must be accountability for the potential effects and consequences of AI systems. Such accountability includes relevant stakeholders to individuals, supervisory authorities and other third parties as appropriate as well as the use of audits, continuous monitoring and impact assessments. The declaration stressed the need to invest in awareness raising, education and research on AI in order to better understand AI and its potential impacts on society and "demonstrable governance processes" for relevant actors, including trusted third parties and independent ethics committees.

**AI systems transparency and intelligibility**

The declaration called for improvements on AI systems' transparency through a variety of means, including investing in public and private scientific research on "explainable" artificial intelligence, making organizational practices more transparent (by promoting algorithmic transparency and the audit-ability of systems and the provision of meaningful information) and ensuring that individuals are always informed appropriately when they are interacting directly with an AI system or when they are providing personal information to be processed by such systems (informational self-determination).

**Ethics by design**

The declaration stressed that AI systems have to be designed and developed responsibly from the very start, applying the principles of privacy by default or privacy by design. Practically, this includes implementing adequate technical and organizational measures and procedures (proportionate to the type of system being designed or implemented) to ensure that data subjects' privacy and personal information are respected. Also, developers should be assessing and documenting the expected or potential impact on individuals and society at large both at the beginning of any AI project and during the project's entire life-cycle and identifying specific requirements for fair and ethical use of such systems.

**Empowerment of individuals**

While the use of AI is to be encouraged, it should not occur at the expense of human rights or the rights of individuals. This includes respecting data protection or privacy rights — including rights to access, the right to object to processing and the right to erasure — and guaranteeing an individual's right not to be subject to a decision based solely on automated processing if the decision significantly impacts them. Regardless,
individuals should always have the right to object or appeal and challenge decisions generated through the use of AI systems.

**Unlawful biases or discrimination**

The declaration expressly acknowledges concerns relating to unlawful bias or discrimination that may occur from the use of data in AI and such unintended results must be reduced and mitigated. Accordingly, developers should invest in research into technical ways to identify, address and mitigate bias, taking reasonable steps to ensure that the personal data or information used in automated decision-making is accurate, up to date and as complete as possible and providing specific guidance and principles in address bias and discrimination, promoting the awareness of individuals and stakeholders.

Recognizing that the declaration’s guiding principles are necessarily pitched at a high level, the ICDPPC also called for the creation of common governance principles on AI to be established on an international basis — given the fact that the development of AI is cross-border activity that will impact everyone.

As part of the declaration, the ICDPPC, therefore, also established a permanent working group — the working group on ethics and data protection in AI — that is now tasked with creating additional guidance to accompany the principles articulated in the declaration.

The group intends to work with all relevant parties involved in the development of AI systems, including governments and public authorities, standardization bodies, AI systems designers, providers, researchers, companies and end users of such systems and will periodically report back to the ICDPPC on its activities.

While there is no doubt that the principles of the declaration are couched in lofty language, they represent a reaction to some very real concerns of individuals that in the rush to seize and capitalize upon the benefits afforded by the use of AI systems — including the ability to process large amounts of data, improving efficiencies and the development of deep-learning technologies — the ethical and human rights must not and cannot be left behind.

Given the landmark work conducted on ethics and AI in this country, spearheaded in Quebec (and further discussed in a prior column) Canada remains well positioned to continue to keep the issues of ethics and data protection at the forefront of AI development.