

Exploring Legal, Ethical and Policy Implications of Artificial Intelligence

DRAFT OCTOBER 2017



WHITE PAPER

EXPLORING LEGAL, ETHICAL AND POLICY IMPLICATIONS OF ARTIFICIAL INTELLIGENCE

By

**Mirjana Stankovic, Ravi Gupta, Bertrand Andre Rossert,
Gordon I. Myers and Marco Nicoli**

September, 2017

Acknowledgments

While it is hazardous to accurately anticipate how the maturation of emerging AI technologies may impact society at large, it is important for policy makers to initiate the process of understanding the truly disruptive nature of these technologies. As AI applications engage in behavior that, were it done by a human, would encroach upon the law, courts and other legal stakeholders would have to puzzle through whom to hold accountable and on what theory. This white paper attempts to lay out central concepts related to AI and complementary technologies, their rapid progress, and discuss the attendant salient legal, ethical, innovation and development policy issues. The paper is not meant to be exhaustive treatment of all issues but rather aimed at providing broad overview of key issues thus stimulating World Bank Group wide discussion on the topic.

This paper was co-authored by Mirjana Stankovic¹, Ravi Gupta², Bertrand A. Rossert³, Gordon I. Myers⁴ and Marco Nicoli⁵.

The paper gratefully benefited from comments provided by Professor Amedeo Santosuosso (ECLT, University of Pavia), Dr. Barbara Bottalico (ECLT, University of Pavia), Professor Oliver Goodenough (Center for Legal Innovation, Vermont Law School), Dr. Bratislav Stankovic (USPTO), Thomas A. Campbell (CEO and President FutureGrasp LLC), and Yupeng Chen (World Bank intern).



This work is licensed under a Creative Commons Attribution 4.0 International License.

¹ Senior Legal Advisor in AI, Industry 4.0, IP, Tech Transfer and Disruptive Technologies, mirjana.stankovic@fulbrightmail.org

² CEO, Tambourine Innovation Ventures, rgupta213@gmail.com; rgupta8@Uchicago.edu

³ Adviser to the Vice President and Chief Ethics Officer, World Bank Group, brossert@worldbank.org

⁴ Chief Counsel, Technology and Private Equity, IFC, gmyers@ifc.org

⁵ Senior Project Manager, Senior Knowledge Management Officer, Global Forum on Law, Justice and Development, World Bank Group, mnicoli@worldbank.org

“By far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.” - Eliezer Yudkowsky.

Contents

Acknowledgments	2
I. Introduction: How AI thinks	5
II. AI and Innovation.....	7
III. AI and Public Policy Concerns.....	12
IV. Implications for Developing Countries	15
V. AI and the Law	19
VI. The Ethics of AI	28
VII. The Way Forward.....	31
Annex 1. AI case law: USA	33
Annex 2 - Bibliography	36
Annex 3 - Who is who.....	39

I. INTRODUCTION: HOW AI THINKS

ARTIFICIAL INTELLIGENCE (AI) IS BROAD CONCEPTUAL TERM FOR TECHNOLOGIES OR SYSTEMS MAKING IT POSSIBLE FOR COMPUTERS TO PERFORM TASKS INVOLVING HUMAN-LIKE DECISION-MAKING, INTELLIGENCE, LEARNED SKILLS AND/OR EXPERTISE.

An accurate understanding of AI competes with its popular portrayal. **While there is no concrete definition of AI, one can attempt to describe it.** AI is a science and a set of computational technologies that are inspired by—but typically operate rather differently from—the ways people use their nervous systems to sense, learn, reason, and take action.

The term was first coined about sixty years ago, when people began trying to understand whether machines can truly think. Since then, there have been advances in search algorithms, machine learning algorithms, and integrating statistical analysis into understanding the world at large over the past six decades.⁶

While concepts related to AI are not new, recent advances have made possible creation of AI tools, and following factors have fueled AI research:

- Rise of the digital economy, which both provides and leverages large amounts of data
- Progress in cloud computing resources, and
- Consumer demand for widespread access to services such as speech recognition and navigation support.

The generic term AI covers a wide range of capabilities. Some futurists such as Stephen Hawking and Sam Harris fear that AI could one day pose an existential threat: a “superintelligence” might pursue goals that prove not to be aligned with the continued existence of humankind. Such fears relate to “strong” AI or “artificial general intelligence” (AGI), which would be the equivalent of human-level awareness, but which does not yet exist. However, such ominous predictions are not shared by other thinkers and practitioners such as Ray Kurzweil, Bill Gates, and Neil de Grasse Tyson.

Current AI applications are forms of “narrow” AI or “artificial specialized intelligence” (ASI). These applications aim to solve specific problems or take actions within a limited set of parameters, some of which may be unknown and must be discovered and learned.⁷ Any time we communicate with a device — book film tickets, pay a gas bill, listen to GPS directions — we

⁶ The History of Artificial Intelligence, History of Computing CSEP 590A, University of Washington, December 2006

⁷ Assessing the risks of Artificial Intelligence, WEF, at <http://reports.weforum.org/global-risks-2017/part-3-emerging-technologies/3-2-assessing-the-risk-of-artificial-intelligence/#view/fn-6> 6

still employ “weak” or “narrow” AI. Consider Apple’s Siri⁸ and Google’s self-drive cars, probably the most recognizable products using “weak” AI. It seems intelligent, but it still has defined functions. It has no self-awareness.⁹

Box 1. AI in our everyday life

Whenever we use our credit card, an AI algorithm approves the transaction. Whenever we use the GPS in our car, we use an AI algorithm. Spam filters are based on AI. The Google translate service is based on statistical machine learning, which is part of AI. The camera face recognition capability of any of our cameras is AI.¹⁰

The essential building blocks of AI are:

Large-scale machine learning which concerns the design of learning algorithms, as well as scaling existing algorithms, in order to work with extremely large data sets. The leap in the performance of algorithms has been accompanied by significant progress in hardware for sensing, perception, and object recognition.

Artificial neural networks (ANNs) are computing systems inspired by the biological neural networks. They learn (progressively improve performance) to do tasks by considering examples, generally without task-specific programming. While impressive, these technologies are highly tailored to particular tasks.

Natural Language Processing is a field that covers computer understanding and manipulation of human language. These technologies analyze human speech for meaning and syntax.

Collaborative systems investigate models and algorithms to help develop autonomous systems that can work collaboratively with other systems and with humans.

Algorithmic game theory and computational social choice draw attention to the economic and social computing dimensions of AI, such as how systems can handle potentially misaligned incentives, including self-interested human participants or firms and the automated AI-based agents representing them.

Reinforcement learning is a framework that shifts the focus of machine learning from pattern recognition to experience-driven sequential decision-making. It promises to carry AI applications forward toward taking actions in the real world.

AI can enable a machine to mimic "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving." Based on deep learning (machines which can

⁸ Siri, the iPhone app that understands us when we speak and responds (usually) in a useful way, is based on AI algorithms for speech understanding.

⁹ Artificial intelligence, Smith, National Magazine (2015), at <http://www.nationalmagazine.ca/Articles/Fall-Issue-2015/Artificial-intelligence.aspx>.

¹⁰ Artificial Intelligence: Potential Benefits and Ethical Considerations, Rossi, European Parliament, at http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/571380/IPOL_BRI%282016%29571380_EN.pdf.

help recognize patterns, speech and natural language) AI is used in mainstream technologies such as web search, medical diagnosis, smart phone applications, and most recently, autonomous vehicles.¹¹

Tasks such as trading stocks, writing sports summaries, flying military planes and keeping a car within its lane on the highway are now all within the domain of ASI. As ASI applications expand, so do the risks of these applications operating in unforeseeable ways, outside the control of humans. The 2010 and 2015 stock market “flash crashes” illustrate how ASI applications can have unanticipated real-world impacts, while AlphaGo¹² shows how ASI can surprise human experts with novel but effective tactics. In combination with robotics, AI applications are already affecting employment and shaping risks related to social inequality.¹³

AI has **great potential to augment** human decision-making by countering cognitive bias and making rapid sense of extremely large data sets. For instance, at least one venture capital firm has already appointed an AI application to help determine its financial decisions. Gradually removing human action can increase efficiency and is necessary for some applications, such as automated vehicles. However, there are **dangers in completely eliminating human oversight** i.e. coming to depend entirely on the decisions of AI systems when we do not fully understand how these systems are making those decisions.¹⁴

II. AI AND INNOVATION

With the advent of autonomous vehicles, digital personal assistants that can anticipate our needs, and computerized health diagnosis, AI is realizing its potential to change people’s lives for the better. The development community is looking at ways to use AI-enabled innovation to achieve key Sustainable Development Goals (SGDs), including low-cost medical diagnosis expert systems, more efficient capital markets, use of automated drones in public emergencies, and many other variants, and to improve the efficiency and effectiveness of their own operations. Thus, by providing new information and improving decision-making through data-driven strategies, AI could potentially help to solve some of the complex global challenges of the 21st century, from climate change and resource utilization to the impact of population growth and healthcare issues.

¹¹ Sunil Johal and Daniel Araya, Work and social policy in the age of artificial intelligence.

¹² <https://en.wikipedia.org/wiki/AlphaGo>

¹³ Supra note 6.

¹⁴ Ibid.

According to a PwC Report “Sizing the price”, global GDP will be 14% higher in 2030 as a result of AI—the equivalent of \$15.7 trillion, more than the current output of China and India combined.¹⁵

Improvements to labor productivity will account for over half of all economic gains from AI between now and 2030, while increased consumer demand resulting from product enhancements will account for the rest.¹⁶

It is estimated that the market for AI services will grow from 420 million U.S. dollars in 2014 to 5 billion U.S. dollars by 2020. Benefits to the broader economy could also be enormous—a recent study by Bank of America Merrill Lynch estimated that AI technology will deliver up to 2 trillion U.S. dollars in cost efficiencies globally. In 2015, over \$2.4 billion in venture capital was invested into the development of AI-based technologies. Well-known companies such as Google, Facebook, Apple and Uber, as well as start-ups, are active in R&D of innovative AI technology-based products. Venture capital funds seeking global AI investment opportunities have become commonplace.

The benefits of AI will be felt differently across sectors. Retail, financial services and healthcare stand to reap the rewards as AI increases productivity, product value and consumption.¹⁷

It is similarly expected that there will be initiatives that explore and implement AI (and its close cousin, machine learning) for use in the public sector, including e-government, anticorruption efforts, and similar activities. The World Bank Group itself is exploring these opportunities.

Every step forward in AI challenges assumptions about what machines can do. A myriad of opportunities for economic benefit have created a stable flow of investment into AI research and development, but with the opportunities come risks to decision-making, security, ethical choices and governance.¹⁸

ROBOTS AND THE MANUFACTURING INDUSTRY (INDUSTRIAL ROBOTS)

Robots already have a demonstrable and significant impact on how manufacturing takes place. Since the start of industrial automation in the 1970s, the use of robots in manufacturing has increased significantly. The industrial robot market was estimated to be worth USD 29 billion in 2014 (including the cost of software, peripherals and systems engineering). The number of robots sold is increasing, reaching about 230,000 units sold in 2014, up from about 70,000 in 1995, and projected to increase rapidly in the next few years. Japan, USA and Europe were the initial leaders in terms of market size. Interestingly, the respective shares of various world regions in

¹⁵ The global economy will be \$16 trillion bigger by 2030 thanks to AI, WEF, at <https://medium.com/world-economic-forum/the-global-economy-will-be-16-trillion-bigger-by-2030-thanks-to-ai-81b37bee96ab>

¹⁶ Ibid.

¹⁷ Id.

¹⁸ Supra note 6.

global robotics sales has changed little, with Asia leading followed by Europe and North America, and rather small volumes in South America and Africa. Yet within Asia, China has gone from no robots in 1995 to overtaking Japan to become the largest robot market. The Republic of Korea is now the second biggest user of industrial robots in Asia.¹⁹

Robots can increase labor productivity, reduce production cost and improve product quality. In the service sector in particular, robots can also enable new business models. Service robots provide assistance to disabled people, mow lawns, but are also deployed in service industries such as restaurants or hospitals. In terms of welfare, robots help humans to avoid strenuous or dangerous work. They also have the potential to contribute solutions to social challenges such as caring for the aging population or achieving environmentally friendly transportation. In part, the economic gains of robots are directly linked to substituting – and thus automating – part of the currently employed workforce.

More productive labor helps keep developed countries' manufacturing firms competitive, avoiding their relocation abroad and creating higher-wage jobs. Many predict that the advancement and proliferation of robotics may lead to an increase in jobs within certain high-income nations as a result of manufacturing re-shoring (in-shoring) because manufacturing previously outsourced to nations with cheaper labor will be performed in high income nations due to robotics reducing the cost of manufacturing. Such predictions are premised on the decrease in manufacturing costs that will result from industrial and other manufacturing related robotics. If such predictions are accurate, it may be true that the proliferation of robotics will increase jobs and economic growth in high income countries. However, at least in the manufacturing sector, the creation of jobs in high income countries may be at the expense of jobs in middle and low-income countries if their primary attraction for their manufacturing industry is low labor costs, further exacerbating income inequality. The use of robots is certain to eliminate both low-skilled but also some types of higher-skilled jobs in the developed countries. Many research studies focus only on the loss of current jobs and do not account for the creation of new job types that may not exist today. Thus, the employment effect of robotics is currently uncertain.

Another question is whether robotic innovation has diffused to developing countries already with meaningful impacts. The installed base of robots outside a few high-income economies and a few exceptions such as China is still limited. However, it is expected that firms involved in manufacturing and assembly activities for global or local supply chains will need to upgrade their use of robots, including some in middle-income or even low-income economies that have so far competed on cheap labor alone. Robots are also gaining ground in low-income countries to address quality issues in local manufacturing.

¹⁹ Breakthrough technologies –Robotics, innovation and intellectual property, Keisner et al, WIPO (2015) at http://www.wipo.int/edocs/pubdocs/en/wipo_pub_econstat_wp_30.pdf; TECHNOLOGY AT WORK The Future of Innovation and Employment, Citi GPS: Global Perspectives & Solutions (2015), at http://www.oxfordmartin.ox.ac.uk/downloads/reports/Citi_GPS_Technology_Work.pdf

AUTONOMOUS VEHICLES

Autonomous vehicles (AV) attract major research spending from car companies as well as internet firms. Their proponents argue that they will reduce road accidents (for instance, through lane-keeping systems, auto-parking, and cruise control), ease congestion, reduce fuel consumption, improve the mobility of the elderly and disabled, and free up commuting time for other tasks. But they also threaten the jobs of millions of people currently employed as drivers. They also raise complex legal issues, such as liability insurance, and intrusion of privacy by hacking.

The European project SARTRE is piloting the concept of “autonomous car platoons,” which allows multiple vehicles to drive autonomously within meters of one another at highway speeds, guided by a professional pilot vehicle. This approach is expected to reduce fuel consumption and emissions by up to 20 percent, improve road safety, and reduce traffic congestion. Drones (unmanned aerial vehicles, and a specialized type of AV) are growing in popularity as prices fall. They have many potential applications, including police work, assisting the disabled, home delivery, farming, entertainment, safety, wildlife conservation, and even providing internet service in remote areas. Rwanda plans to be home to the world’s first drone airport, or “droneport,” to facilitate the delivery of medical and emergency supplies, quickly and cost effectively, across geographical barriers. The introduction of AVs is likely to be gradual, with many cars and planes already incorporating elements of assistive technology. The impact on jobs will ultimately be a function of price (self-driving cars are currently prohibitively expensive), legislation (will they always require a human with manual override?), and time.²⁰

Box 2. Economic disruption caused by AV

One of the biggest unresolved policy questions is how to deal with the economic disruption that would affect millions of taxi drivers, chauffeurs and truck drivers whose jobs could potentially be eliminated by self-driving vehicles. The first self-driving truck began testing in the deserts of Nevada in May 2015, and firms such as Daimler and Otto (a startup launched by two former Google engineers) are now attempting to perfect the technology. Morgan Stanley predicts completely autonomous capability by 2022 and massive market penetration by 2026.

While autonomous trucks could surely help reduce large-truck crashes, they could also eliminate the jobs of millions of truckers and millions of non-drivers employed within the trucking industry. The technology could also threaten the jobs of millions of people who work in restaurants, motels and truck stops that service truck drivers, with community ripple effects flowing from those job losses. The appropriate policy responses to such developments — re-training? a basic income? —

²⁰ Digital Dividends, World Development Report, World Bank (2016), at

<http://documents.worldbank.org/curated/en/896971468194972881/pdf/102725-PUB-Replacement-PUBLIC.pdf>

are relatively unexplored.²¹

THE INTERNET OF THINGS

The “internet of things” (IoT) refers to the interconnection of objects to internet infrastructure through embedded computing devices, such as radio frequency identification (RFID) chips and sensors. IoT products can be classified into five broad categories: wearable devices, smart homes, smart cities, environmental sensors, and business applications. Cisco estimates that by 2020, 50 billion devices and objects will be connected to the internet. IoT is quickly redefining service delivery and unlocking opportunities in multiple areas. Smart fitness sensors and trackers are transforming health care and improving personal fitness and health. Embedded sensors accurately relay moisture, air and water pollution levels, and resource levels, allowing for closer monitoring of environmental problems. Factories and supply chains use smart sensors to improve the efficiency of manufacturing and distribution of goods. Globally, there has been a rise in spaces where people can gather to build and learn with electronics, software, and digital fabrication. Known as makerspaces, these spaces have democratized access to tools and empowered participants to build and learn on their own. One of the key applications of IoT is in combating climate change and its effects. Farms in developing countries can use intelligent sensors to monitor soil conditions and guide autonomous irrigation systems. Smart traffic synchronization systems in cities save on travel time and fuel consumption. Countries such as Singapore are deploying smart networks that use global positioning systems (GPS), sensor information from monitoring cameras, and other sources to sense population movement, ease traffic congestion, and re-route traffic in the case of special events and emergencies. Some experts believe that the IoT will mark a new stage of the internet’s development, since it has the potential to revolutionize the way people live, work, interact, and learn. However, there are still significant barriers to full commercialization of IoT, such as the fragmented landscape of standardization, which is preventing interoperability; and the relatively high cost of embedded devices. The maker movement offers a possible solution for the standardization challenge, empowering individuals to adjust devices to fit the local context. There are also significant privacy and security concerns. As more devices are connected to networks, hacking unsecure devices could have repercussions that far exceed the damage posed by conventional security threats.²²

EMPLOYMENT

The hierarchy of labor is concerned primarily with automation. As humans have invented ways to automate jobs, AI could create room for people to assume more complex roles, moving from the

²¹ Artificial Intelligence Comes of Age, The Promise and Challenge of Integrating AI into Cars, Healthcare and Journalism, Bollier, The Aspen Institute (2017), at <http://csreports.aspeninstitute.org/Roundtable-on-Artificial-Intelligence/2016/report/details/0200/AI-2016>

²² Supra note 20.

physical work that dominated the pre-industrial globe to the cognitive labor that characterizes strategic and administrative work in our globalized society.

For instance, trucking currently employs millions of individuals in the United States alone. What will happen to them if the self-driving trucks promised by Tesla's Elon Musk become widely available in the next decade? But on the other hand, if we consider the possibility for lower risk of accidents, self-driving trucks might be an ethical choice. The same scenario could happen to office workers, as well as to the majority of the workforce in developed countries.

Current AI developments have not yet impacted hugely white-collar jobs. AI's impact is largely felt in the blue-collar, repetitive jobs segment. Those who are inventing the technologies can play an important role in easing the effects of AI on jobs. The solution is not to hold back on innovation, but we have a new problem to innovate around: how do you keep people engaged when AI can do most things better than most people? When governments choose what research to fund and when businesses decide what technologies to use, they are inevitably influencing jobs and income distribution. The predicament lies in the fact that it is not easy to see a practical mechanism for picking technologies that favor a future in which more people have better jobs.²³

AI START UPS

Start-ups specializing in AI applications received US\$2.4 billion in venture capital funding globally in 2015 and more than US\$1.5 billion in the first half of 2016.² Government programs and existing technology companies add further billions. Leading players are not just hiring from universities, they are hiring the universities: Amazon, Google and Microsoft have moved to funding professorships and directly acquiring university researchers in the search for competitive advantage.²⁴

III. AI AND PUBLIC POLICY CONCERNS

INEQUALITY. HOW DO WE DISTRIBUTE WEALTH CREATED BY AI?

Today's economic systems are based on the principle of compensation for contribution to the economic unit, which is assessed through hourly wage. AI is a game changer in this respect because by using AI, a company can drastically cut down on relying on the human workforce. Thus, the earnings will go to fewer people. Consequently, individuals who have ownership in AI-driven companies would benefit disproportionately at the expense of workforce which was eliminated as a result of AI.

²³ Who will own the robots, Rotman, MIT Technology Review (2015), at <https://www.technologyreview.com/s/538401/who-will-own-the-robots/>

²⁴ Supra note 6.

We are already seeing a widening wealth gap, where start-up founders take home an increasingly larger portion of the economic surplus they create. In 2014, roughly the same revenues were generated by the three biggest companies in Detroit and the three biggest companies in Silicon Valley. However, in Silicon Valley there were 10 times fewer employees for the comparable firms.

One question remains to be answered by public policy stakeholders: How does a fair society achieve an equitable post-labor economy?²⁵

HUMANITY. HOW DOES AI AFFECT OUR BEHAVIOR AND INTERACTION?

Artificially intelligent bots are becoming better at modelling and mimicking human behavior. In 2015, a bot named Eugene Goostman won the Turing Challenge for the first time. In this challenge, human raters used text input to chat with an unknown entity, then guessed whether they had been chatting with a human or a machine. Eugene Goostman fooled more than half of the human raters into thinking they had been talking to a human being.

This milestone is only the start of an age where we will frequently interact with machines as if they are humans; whether in customer service or sales. While humans are limited in the attention and kindness that they can expend on another person, artificial bots can channel virtually unlimited resources into building relationships.

We are already witnessing how machines can trigger the human brain's reward (e.g. click-bait headlines and video games). These headlines are often optimized with a rudimentary form of algorithmic optimization for content to capture our attention. These methods are used to make numerous video and mobile games become addictive.²⁶

ARTIFICIAL STUPIDITY

Systems usually have a training phase in which they "learn" to detect the right patterns and act according to their input. Once a system is fully trained, it can then go into test phase, where it is hit with more examples and we see how it performs. The training phase cannot cover all possible examples that a system may deal with in the real world. These systems can be fooled in ways that humans wouldn't be.²⁷ For instance, random dot patterns can lead a machine to "see" things that are not there. If the world wants to reap of the benefits offered by AI, there should be a governance system with checks and balances to ensure that the machine performs as intended,

²⁵ Top 9 ethical issues in artificial intelligence, Bossmann, WEF (2016), at <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/>.

²⁶ Why your favorite apps are designed to addict you, Coca, The Kernel, at <http://kernelmag.dailydot.com/issue-sections/features-issue-sections/15708/addicting-apps-mobile-technology-health/>.

²⁷ Deep neural networks are easily fooled: High confidence predictions for unrecognizable images, Evolving Artificial Intelligence Laboratory, University of Wyoming, at <http://www.evolvingai.org/fooling>.

and that humans cannot trick it to use it for their own ends which may sometimes be reprehensible.

AI BIAS

Though AI is capable of a speed and capacity of processing that's far beyond that of humans, it cannot always be trusted to be fair and neutral. Google and its parent company Alphabet are one of the leaders when it comes to artificial intelligence, as seen in Google's Photos service, where AI is used to identify people, objects and scenes. But it can go wrong, such as when a camera missed the mark on racial sensitivity²⁸, or when a software used to predict future criminals showed bias against black people.²⁹ Clearly the AI system was internalizing the prejudices of its creators who had programmed the AI thus inadvertently passing on its own biases.

AI systems are created by humans, who can be biased and judgmental. Once again, if used right, or if used by those who strive for social progress, artificial intelligence can become a catalyst for positive change.³⁰

SECURITY

The more powerful a technology becomes, the more potently can it be used for reprehensible reasons as well as good. This applies not only to robots produced to replace human soldiers, or autonomous weapons, but to AI systems that can cause damage if used maliciously. Consequently, cybersecurity will become even more important.

Box 3. Internet of Bodies

For years, cybersecurity has focused on physical objects that comprise the Internet of Things, but with technological advances in healthcare, this now includes medical implants. What began with external, smart objects like FitBits, has steadily grown to internet-connected pacemakers, cochlear and microchip implants, and more. We now face a new era of imperative and legal security research of consumer devices that are attached to both the Internet and the human body. This "Internet of Bodies" will inevitably expose us to unprecedented cybersecurity vulnerabilities, introduce conflict across several legal regimes, and raise fundamental ethical questions about the future of what it means to be human in an age of technology-mediated bodies and artificial intelligence.³¹

²⁸ Camera misses the mark on racial sensitivity, Lee, Gizmodo (2009), at <http://gizmodo.com/5256650/camera-misses-the-mark-on-racial-sensitivity>.

²⁹ Machine Bias There's software used across the country to predict future criminals. And it's biased against blacks, Angwin et al., Propublica (2016), at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

³⁰ See supra note 25.

³¹ Atlantic Council Cyber Statecraft Initiative, Dr. Andrea Matwyshyn's project "Internet of Bodies".

EVIL AI

There is a possibility of turning AI to fulfill human wishes with terrible unforeseen consequences. Or there is a possibility of AI's lack of understanding of the full context in which the wish was made. Imagine an AI system that is asked to eradicate cancer in the world. After a lot of computing, AI finds a formula that will eradicate cancer – by killing everyone on the planet!

SINGULARITY. HOW DO HUMANS STAY IN CONTROL OF A COMPLEX INTELLIGENT SYSTEM?

Human dominance is entirely due to the cognitive revolution that was triggered by human ability to imagine things, develop tools and intelligently design the environment.

This poses a serious question about artificial intelligence: will it, one day, have the same advantage over us? We cannot rely on just "pulling the plug" either, because a sufficiently advanced machine may anticipate this move and defend itself. This is what some call the "singularity": the point in time when human beings are no longer the most intelligent beings on earth.³²

AI RIGHTS

Humans develop mechanisms of reward and aversion in AI systems. For instance, in the case of reinforcement learning improved performance is reinforced with a virtual reward, just like training a dog.

Currently AI systems are still superficial, however in near future they might become life-like. If the AI system is "punished" with negative input, will it suffer? If AI is enabled to perceive, act and feel, will it be awarded legal status? Should AI be granted certain rights analog to the rights of corporations?³³

IV. IMPLICATIONS FOR DEVELOPING COUNTRIES

The Accenture Institute for High Performance recently released research revealing that, by 2035, AI could double annual economic growth rates in developed economies. The study compared economic output in each country in 2035 under a baseline scenario based on current assumptions against one showing expected growth once the impact of AI has been absorbed into the economy. In the US, the annual growth rate went up from 2.6% to 4.6% - an additional USD \$8.3 trillion in gross value added (GVA) with widespread AI adoption included. In the UK, AI could add an

³² See supra note 25.

³³ Ibid.

additional USD \$814 billion to the economy, increasing the annual growth rate of GVA from 2.5% to 3.9%.³⁴

However, it is developing economies where AI is likely to have significant impact. We have already entered a period in which enormous technology-driven change is helping to address a number of challenges in developing economics. AI technology, in particular, has extremely strong developmental implications. The rate at which new knowledge is transferred to less advanced economies for the first time, have been converging over time across countries (Comin and Ferrer 2013). There are still a number of challenges towards implementation of such technology. The infrastructure, for one, is not necessarily capable of integrating all AI technologies, so it is not a case of inventing something new and then dropping it into a developing economy. Yet, while the priority must be to build infrastructure - next-generation telecoms, power and agriculture systems - so that AI can be used, there are already a number of ways that it can be applied.³⁵ The introduction of AI could also widen the digital divide. Without a computer connected to internet, many everyday activities are no longer possible. Similarly, access to AI benefits may be inadvertently limited to advantaged populations, providing them with exponentially superior knowledge and tools, and further separating the poor from equitable participation in the global economy. Penetration rates of new technologies within countries and their adoption by enterprises to upgrade productivity have been diverging—with large gaps within countries between the best and average technological practices within each industry (Comin and Ferrer 2013). Low technology diffusion is increasingly recognized as a driver for exacerbating inequality (Dutz 2014, Piketty 2014) and recent literature points to need for governments to tilt their policies to address this challenge.

The challenge of assuring access and inclusion to AI is indeed a core practitioner and policy concern.

AGRICULTURE

There are two pressing concerns for the majority of people in developing countries: access to water and food. To provide citizens with food, smallholder farms must be able to produce enough. However, currently, research infrastructure and agricultural extension systems capable of supporting smallholder farmers are sadly lacking. AI is capable of increasing the yield of farmland under tillage in developing countries, with machine learning algorithms used in drone technology to both plant and fertilize seeds at a speed beyond human abilities.

Another application of AI for food management in developing economics is identification of disease in crops so they can be more easily treated. A team of researchers at Penn State and the

³⁴ Artificial Intelligence is the Future of Growth, Accenture, at <https://www.accenture.com/us-en/insight-artificial-intelligence-future-growth>.

³⁵ AI In Developing Countries, Artificial intelligence isn't just for self driving cars, at <https://channels.theinnovationenterprise.com/articles/ai-in-developing-countries>.

Swiss Federal Institute of Technology (EPFL) have fed a network of computers with over 53,000 photos of both healthy and unhealthy plants in an attempt to recognize specific plant diseases. Such technology will provide the basis for field-based crop-disease identification using smartphones. The system has been able to identify both crops and diseases – from photos – with an accuracy rate of up to 99.35%.³⁶

Numerous multiple variable optimization problems that have eluded solutions may become tractable with AI. For instance, AI has shown promise to untangle the vexed water energy nexus issue. The two systems (water and energy) are currently optimized but to each other's detriment. With increasing stress on resources, a system where both resources are optimized is direly needed.³⁷

For NGOs and charities, determining where resources are needed is vital to helping those most in need. If available resources are not properly utilized, the scarcity makes a bigger dent. This is another area where AI can help greatly. AI can be used to learn to analyze multiple factors at the same time in a way that humans cannot which can show, say, where a drought could occur, how many people it is likely to impact, and what is required to fix the problem.

For example, 'Harvesting' is a startup machine learning to analyze satellite data of the Earth's surface. The machine is trying to pinpoint areas in need of investment in water and tools needed for farming to help institutions distribute money more efficiently.

CEO of machine learning startup Harvesting, Ruchit Garg, recently noted of AI that, 'Our hope is that in using this technology we would be able to segregate such farmers and villages and have banks or governments move dollars to the right set of people.'³⁸

HEALTHCARE

The Ebola virus wreaked havoc on African communities, as numerous outbreaks have over the years. In the case of Ebola, Barbara Han, a disease ecologist at the Cary Institute of Ecosystem Studies, said, 'Using machine learning methods developed for artificial intelligence, we were able to bring together data from ecology, biogeography, and public health to identify bat species with a high probability of harboring Ebola and other filoviruses. Understanding which species carry these viruses, and where they are located, is essential to preventing future spillovers.'

The main advantage of Machine Learning is its ability to deal with complexities. With a number of variables interacting at one time, findings can become difficult to interpret. Machine Learning side steps this. On this issue Han says: 'The algorithm doesn't care how the variables are interacting; its only goal is to maximize predictive performance. Then we human scientists can step up.'

³⁶ Ibid.

³⁷ Id.

³⁸ Id.

Machine learning technology is the most effective way of not only understanding the spread of disease, but also providing relief. We are looking at future where machine learning could feasibly identify a disease, develop a cure, locate where the outbreak is likely to strike next, and then transport the cure there in autonomous vehicles, all with minimal human interaction.³⁹

OPEN SOURCE INITIATIVES

Robotics platforms used in universities and businesses are increasingly central to robotics innovation. Often these platforms are based on open-source software such as the Robot Operation System (ROS). These open source robotics platforms invite third parties to use and/or improve existing content without the formal negotiation or registration of IP rights. Instead, software or designs are distributed under Creative Commons or GNU General Public License, a free software license. This allows for rapid prototyping and flexible experimentation. Various non-profit organizations and projects support the development, distribution and adoption of open-source software for use in robotics research, education and product development. The iCub⁴⁰, for instance, is an open-source cognitive humanoid robotics platform funded by the EU which has been adopted by a significant number of laboratories. Poppy is an open-source platform developed by INRIA Bordeaux for the creation, use and sharing of interactive 3D-printed robots.⁴¹ Other examples include the Dronecode project (open source, collaborative project that brings together existing and future open source drone projects under a nonprofit structure governed by The Linux Foundation. The result will be a common, shared open source platform for Unmanned Aerial Vehicles) and the NASA International Space Apps Challenge⁴².

Some of this will entail an increasing shift toward engaging end-users or amateur scientists to interact and improve on existing robotics applications. In fact, many user-oriented low-cost platforms built for home or classroom use, like TurtleBot and LEGO Mindstorms, are built on open-source platforms. This open-platform approach is not limited to software; it can also encompass blueprints such as technical drawings and schematics, including designs. The Robotic Open Platform (ROP), for instance, aims to make hardware designs of robots available to the robotic community under an Open Hardware license; advances are shared within the community.

While open source helps practitioners rapidly build on existing knowledge it also raises vexed issues on assignment of liability should an AI application built with such diffuse ownership structure go awry.

³⁹ Id.

⁴⁰ <http://www.icub.org>

⁴¹ <https://www.poppy-project.org/>

⁴² <https://2017.spaceappschallenge.org/>

V. AI AND THE LAW

Emerging AI is an ever-increasing public concern for the many risks present where decisions are made by computers and not by humans. AI requires access to vast amounts of data, but poorly drawn laws and government policies can hinder beneficial access without reducing the risk of AI activities. AI also raises important ethical and privacy concerns that could erode trust in emerging technologies if not addressed thoughtfully.

As a transformative technology, AI has the potential to challenge numerous legal assumptions in the short, medium, and long term. Precisely how law and policy will adapt to advances in AI—and how AI will adapt to values reflected in law and policy—depends on a variety of social, cultural, economic, and other factors, and is likely to vary by jurisdiction. A common thread, however is that policymakers will be challenged to draft legislation that does not stifle AI innovation, but at same time protects the public from possible dangers presented when computer judgment replaces that of humans.

Since policy formulation is rarely swift, the judiciary may be the first to address these novel legal issues. As a result, practitioners will bear a special responsibility to understand the emerging issues, interests and arguments involved in AI activities, in order to guide their clients.

While we do not intend to conduct comprehensive examination of the ways AI interacts with the law, our objective is to foster a discussion on key legal and policy issues related to AI.

MODERNIZE LAWS AND PRACTICES TO ENABLE AI

AI requires access to data—machines cannot “learn” unless they have large data sets from which to discern patterns. Governments should carefully assess whether existing data access laws should be updated to reflect the benefits of AI. For example, while copyright laws should protect the expressive value of a work, policymakers may wish to extend “fair use” or similar concepts to allow AI insights in ways that do not compete with copyright owners. When it comes to personal information, governments should consider appropriate consumer protection and privacy laws, including “safe harbor” masking technologies, opt-in rights, data minimization requirements, and similar policies that appropriately balance privacy concerns against the benefits of AI insights based on access to data.

Governments can also leverage the transformational impact of cloud computing by encouraging legal frameworks for establishing common data pools for AI innovations shared in ways that do not disclose personal data, trade secrets or other proprietary information, or encourage anticompetitive activity. In addition, governments should ensure that all data they collect is available to the public for analysis, subject to privacy and national security considerations.

INTELLECTUAL PROPERTY (IP) LAW

IP laws are key to much of the recent push for AI-related legislation. For instance, with the exception of the Copyright Act, the US IP law, including the Intellectual Property Clause of the Constitution, do not explicitly mention that in order for creative works or novel inventions to be protected by law, they must be the result of human efforts. Similarly, the World Intellectual Property Organization’s definition of intellectual property “refers to creations of the mind,” but does not explicitly require that the “mind” be human. Assuming nonetheless that public policy will want to reward and protect individual creators, there may be discussion whether the policy is best served by presumptively viewing the AI inventor or user as the “author” of an AI-generated work (or derivative work).

Regarding patent law, the traditional aims of incentivizing both invention and public disclosure by way of limited monopoly, may require rethinking who the “inventor” is, where the invention is patentable subject matter generated by an AI function.

Policymakers may wish to support and incentivize open source approaches to AI development, in order to avoid the anticompetitive tendencies that some argue are inherent in platform technologies. They may also wish to consider the extent “business method” patents should be granted, consistent with public policy aims; and to the extent allowed, how examiners can avoid allowing overbroad claims that for example, inadvertently extend the statutory patent monopoly to AI-enabled inventions otherwise unrelated to the original grant.

Patent Litigation: US law perspective

The main issue courts have addressed in US patent litigation cases is whether the AI subject matter is patent-eligible subject matter under 35 U.S.C. § 101. Courts addressing this question must first ask whether a patent’s claims are directed to a patent-ineligible concept, such as laws of nature or abstract ideas. If not directed to such a concept, a patent will be enforceable under this test. However, if a patent’s claims are directed to a patent-ineligible concept, the analysis moves to a second step: whether the patent claims, despite being directed to a patent-ineligible concept, are nevertheless patent-eligible because they include a sufficiently “inventive concept”—an element or combination of elements that is sufficient to ensure that the patent in practice amounts to significantly more than a patent upon the ineligible concept itself.⁴³

⁴³ See *Vehicle Intelligence & Safety LLC v. Mercedes-Benz USA, LLC*, 635 F. App’x 917 (Fed. Cir. 2015), cert. denied, 136 S. Ct. 2390 (2016), (dismissing certain claims directed to the use of “expert system(s)” to screen equipment operators for impairments such as intoxication as patent-ineligible). The Vehicle Intelligence Court first determined that the claims at issue were directed to a patent-ineligible concept—“the abstract idea of testing operators of any kind of moving equipment for any kind of physical or mental impairment.” The “expert system” concept was considered abstract because, based on the definition assigned to it by the Court during claim construction, it was something performed by humans absent automation, and also because “neither the claims at issue nor the specification provide any details as to how this ‘expert system’ works or how it produces faster, more accurate and reliable results.” This lack of clarity contributed to a holding of lack of inventive concept in the second step, rendering the patent claims at issue

At least one US District Court opinion has considered the patentability of driverless cars and automated support programs. In *Hewlett Packard Co. v. ServiceNow, Inc.*, No. 14-CV-00570-BLF, 2015 WL 1133244 (N.D. Cal. Mar. 10, 2015), Judge Freeman of the Northern District of California while a self-driving car may be very commercially successful, novel, and non-obvious, the concept of a self-driving car is still abstract. While an inventor “may be able to patent his specific implementation,” Judge Freeman disagreed that the concept of self-driving cars could be patented in the abstract. While Judge Freeman’s hypothetical is likely dicta, it nevertheless serves as a guidepost regarding patent eligibility of self-driving vehicles.

AI issues also occur in the area of determining the issue of inventorship. According to US patent law, an inventor can use “the services, ideas, and aid of others in the process of perfecting his invention without losing his right to a patent.”⁴⁴ Also, “patentability shall not be negated by the manner in which the invention was made.”⁴⁵ However, the patent law defines “inventor” as “the individual . . . who invented or discovered the subject matter of the invention” and the statutes also describe joint inventors as the “two or more persons” who conceived of the invention. See 35 U.S.C §§ 100, 116(a). The Federal Circuit has explicitly barred legal entities from obtaining inventorship status because “people conceive, not companies.” *New Idea Farm. Equip. Corp. v. Sperry Corp.*, 916 F.2d 1561, 1566 n.4 (Fed. Cir. 1990).⁴⁶

Copyright law: US law perspective

The US Copyright Office has announced that it “will not register works produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author.”⁴⁷

unenforceable. The Federal Circuit compared the patent as equivalent to “a police officer field-testing a driver for sobriety.” In *Blue Spike, LLC v. Google Inc.*, No. 14-CV-01650-YGR, 2015 WL 5260506, at *5 (N.D. Cal. Sept. 8, 2015), *aff’d*, 2016 WL 5956746 (Fed. Cir. Oct. 14, 2016), the Court found that because the patents at issue sought to model on a computer “the highly effective ability of humans to identify and recognize a signal,” the patents simply cover a general-purpose computer implementation of “an abstract idea long undertaken within the human mind.” The Blue Spike Court also found that the second step of the eligibility inquiry for “inventive concept” was not present as the claims “cover a wide range of comparisons that humans can, and indeed, have undertaken since time immemorial.”

⁴⁴ *Hess v. Advanced Cardiovascular Sys.*, 106 F.3d 976, 981 (Fed. Cir. 1997).

⁴⁵ 35 U.S.C. Section 103.

⁴⁶ Artificial Intelligence Litigation: Can the Law Keep Pace with The Rise of the Machines?, at <http://www.quinnemanuel.com/the-firm/news-events/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/>.

⁴⁷ U.S. Copyright Office, *The Compendium of U.S. Copyright Office Practices* § 306 (3d ed. 2014); see also U.S. Copyright Office, *The Compendium of U.S. Copyright Office Practices* § 202.02(b) (2d ed. 1984), available at <http://copyright.gov/history/comp/compendium-two.pdf> (“The term ‘authorship’ implies that, for a work to be copyrightable, it must owe its origin to a human being.”) The 2014 iteration of the Human Authorship Requirement was partially the result of a prominent public discourse about non-human authorship stemming from the “Monkey

CIVIL (CONTRACT AND TORT) LIABILITY

The more autonomous AI applications are, the more difficult it may be to hold individual authors or inventors liable for the (increasingly less foreseeable) consequences of AI activities. This makes the ordinary rules on liability insufficient and calls for new rules which focus on how to determine whether an AI application is responsible for its acts or omission; and if so, whether policy considerations nonetheless suggest that the vendor or the inventor should bear the resulting cost. If the purpose is to incentivize due care, a strict liability regime backed by insurance may be most efficient.

Beyond these immediate concerns, as autonomous AI applications develop, the fundamental question of whether AI should possess a legal status may be addressed by legislators worldwide. The issue of AI autonomy would raise the question of its nature in the light of the existing legal categories – of whether it should be regarded as natural person, legal person, animal or object – or whether a new category should be created, with its own specific features and implications as regards the attribution of rights and duties, including liability for damage.

Unlike legislation, the protection provided by the courts is remedial not preventative. Courts assess liability and damages based on prior legal precedent. Cases where the harm is alleged to have been caused by AI applications ask the court to unravel novel technology and apply ill-fitting case law to make determinations of liability. For example, US common law tort and malpractice claims often center on human centered concepts of fault, negligence, knowledge, intent, and reasonableness. What happens when human reasoning is replaced by a AI application? What happens when the perpetrator or the victim is AI? ⁴⁸

Claims regarding AI are novel and there is no well-established jurisprudence in this regard. US common law claims involving analogous automated technology can be used as an analytical framework in this regard (See *In re Ashley Madison Customer Data Sec. Breach Litig.*, 148 F.

Selfies.” See *Naruto v. Slater*, No. 3:2015-cv-04324, 2016 WL 362231, *1 (N.D. Cal. Jan. 23, 2016). While there have not yet been cases tackling this unique issue of inventorship, scholars have begun to take notice and weigh in. See, e.g., Ben Hattenbach, Joshua Glucoft, Patents in an Era of Infinite Monkeys and Artificial Intelligence, 19 *Stan. Tech. L. Rev.* 32 (2015); Ryan Abbott, I Think, Therefore I Invent: Creative Computers and the Future of Patent Law, 57 *B.C. L. Rev.* 1079 (2016).

⁴⁸ Artificial Intelligence Litigation: Can the Law Keep Pace with The Rise of the Machines?, at <http://www.quinnemanuel.com/the-firm/news-events/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/>.

Supp. 3d 1378, 1380⁴⁹; *Go2Net, Inc. v. C I Host, Inc.*, 115 Wash. App. 73 (2003)⁵⁰; Cases involving personal injury resulting from automated machines have also been litigated⁵¹).

Reinforcement learning

Unlike other methods of generating AI models, reinforcement learning is a training method that allows AI models to learn from its past experiences. With reinforcement learning AI must determine the best course of action in order to get a high score. In near future, reinforcement learning AI would be integrated with more hardware and software solutions, such as AI-controlled traffic signals capable of adjusting light timing to optimize the flow of traffic or AI-controlled drones capable of optimizing motor revolutions to stabilize videos.

As reinforcement learning gains more independent decision-making capabilities, the law will have to adjust to the new reality. For instance, what will happen if the AI-controlled traffic signal learns that it is most efficient to change the light one second earlier than previously done, but that causes more drivers to run the light and causes more accidents?

⁴⁹ A decision in a consolidated class action in the District Court for the Eastern District of Missouri found that the use of a computer program to simulate human interaction could give rise to liability for fraud. Among the claims related to a data breach on the infamous Ashley Madison online dating website in 2015 that resulted in mass dissemination of user information, were allegations that defendants were engaging in deceptive and fraudulent conduct by creating fake computer “hosts” or “bots,” which were programmed to generate and send messages to male members under the guise that they were real women, and inducing users to make purchases on the website. It is estimated that as many as 80% of initial purchases on the website—millions of individual transactions—were conducted by a user communicating with a bot operating as part of Ashley Madison’s automated sales force for the website.

⁵⁰ Another court, in a case involving an internet advertising breach of contract claim, was asked to resolve a dispute over the meaning of “impressions,” a key term in Internet advertising. The Go2Net Court determined that the parties’ contract permitted visits by search engines and other “artificial intelligence” agents, as well as human viewers, in the advertiser’s count of “impressions.”

⁵¹ For example, cases have involved workers compensation claims or claims against manufacturers by workers injured by robots on the job. See, e.g., *Payne v. ABB Flexible Automation, Inc.*, 116 F.3d 480, No. 96-2248, 1997 WL 311586, *1-*2 (8th Cir. 1997) (per curiam) (unpublished table decision); *Hills v. Fanuc Robotics Am., Inc.*, No. 04-2659, 2010 WL 890223, *1, *4 (E.D. La. 2010); *Bynum v. ESAB Grp., Inc.*, 651 N.W.2d 383, 384-85 (Mich. 2002) (per curiam); *Owens v. Water Gremlin Co.*, 605 N.W.2d 733 (Minn. 2000). There has also been extensive litigation over the safety of surgical robots, especially the “da Vinci” robot manufactured by Intuitive Surgical, Inc. See, e.g., *O'Brien v. Intuitive Surgical, Inc.*, No. 10 C 3005, 2011 WL 304079, at *1 (N.D. Ill. Jul. 25, 2011); *Mracek v. Bryn Mawr Hosp.*, 610 F. Supp. 2d 401, 402 (E.D. Pa. 2009), aff'd, 363 F. App'x 925 (3d Cir. 2010); *Greenway v. St. Joseph's Hosp.*, No. 03-CA-011667 (Fla. Cir. Ct. 2003). Although the court in *United States v. Athlone Indus., Inc.*, 746 F.2d 977, id. at 979 (3d Cir. 1984) stated that “robots cannot be sued” and discussed instead how the manufacturer of a defective robotic pitching machine is liable for civil penalties for the machine's defects, it is important to note that this decision was rendered in 1984.

The traditional concept of civil (contract and tort) liability might be less easily applied to developments in autonomous AI, particularly in a scenario where a AI application might cause damage that cannot be easily or foreseeably traced back to human error. Many aspects of civil liability law might, then, need rethinking, including basic civil liability law, accountability for damage, or its social relevance. It may become difficult to ascertain what caused the damage in certain situations, particularly if AI is able to learn new things by itself.

The current law traditionally finds liability where the developer was negligent or could foresee harm. For example, the court in the US case *Jones v. W + M Automation*⁵², Inc., a case from New York state in 2007, did not find the defendant liable where a robotic gantry loading system injured a worker, because the court found that the manufacturer had complied with regulations. Typically, in any US state, a plaintiff asserting a strict liability claim against a robot manufacturer must plead and prove that the defendant sold a product that was defective and unreasonably dangerous at the time it left the defendant's hands, the product reached the plaintiff without substantial change, and the defect was the proximate cause of plaintiff's injuries. Under a negligent design theory, a plaintiff would seek to show a robotics manufacturer had a duty to exercise reasonable care in manufacturing the robot, the manufacturer failed to exercise reasonable care in making the robot, and the defendant's conduct proximately caused plaintiff's damages.⁵³

The challenge with reinforcement learning is that there is no fault by humans and no foreseeability of an encroachment. Under the traditional tort law, the AI developer would not be held liable. The main question that remains to be answered is whether the law will adapt to the new technological reality so that the world enters a dystopian future where AI is held responsible for its own actions and given personhood. One way to go forward is the soft governance approach by adopting AI ethical standards mandated through international law where manufacturers and developers agree to abide by general ethical guidelines. This should be done through a democratic process of convening key AI stakeholders, such as governments, international organizations, expert groups (e.g. OpenAI⁵⁴), private sector, NGOs, consumers, and establishing a standard that includes explicit definitions for neural network architectures, as well as quality standards to which AI must adhere.⁵⁵

Box 4. EU rules on civil liability in light of robot autonomy

(i) Under the current legal framework robots cannot be held liable per se for acts or omissions that cause damage to third parties. Also, the existing rules on liability cover cases where the cause of the robot's act or omission can be traced back to a specific human agent such as the

⁵² Summary of selected robotics cases, Stephen S. Wu, at http://ftp.documation.com/references/ABA10a/PDFs/2_5.pdf.

⁵³ Ibid.

⁵⁴ OpenAI is a non-profit AI research company, discovering and enacting the path to safe artificial general intelligence at <https://www.openai.com/>

⁵⁵ Artificial intelligence and the law, Elman and Castilla, TechCrunch, <https://techcrunch.com/2017/01/28/artificial-intelligence-and-the-law/>.

manufacturer, the owner or the user and where that agent could have foreseen and avoided the robot's harmful behavior.

(ii) Manufacturers, owners or users could be held strictly liable for acts or omissions of a robot if, for example, the robot was categorized as a dangerous object or if it fell within product liability rules. In the scenario where a robot can make autonomous decisions, the traditional rules will not suffice to activate a robot's liability, since they would not make it possible to identify the party responsible for providing compensation and to require this party to make good the damage it has caused.

(iii) Machines designed to choose their counterparts, negotiate contractual terms, conclude contracts and decide whether and how to implement them make the traditional contractual liability rules inapplicable. Regarding non-contractual liability, Council Directive 85/374/EEC of 25 July 1985 (Liability for defective products) can only cover damage caused by a robot's manufacturing defects and on condition that the injured person is able to prove the actual damage, the defect in the product and the causal relationship between damage and defect (strict liability or liability without fault). Notwithstanding the scope of the Liability for defective products Directive, the current legal framework would not be sufficient to cover the damage caused by the new generation of robots, as they can be equipped with adaptive and learning abilities entailing a certain degree of unpredictability in their behavior, since these robots would autonomously learn from their own, variable experience and interact with their environment in a unique and unforeseeable manner.

The European civil rules in robotics propose that “*the future legislative instrument should provide for the application of strict liability as a rule, thus requiring only proof that damage has occurred and the establishment of a causal link between the harmful behavior of the robot and the damage suffered by the injured party*”. This would, then, be a strict liability regime, which could be labelled “*vicarious liability for the robot(s)*.” Here the double burden of proof falls to the victim of the damage. Yet, even in this specific case, deciding who is the ultimate respondent, i.e. where responsibility truly lies, would remain tricky.”⁵⁶

Autonomous vehicles

As autonomous vehicles become more widespread, questions will arise over their security, including how to ensure that technologies are safe and properly tested under different road conditions prior to their release. Autonomous vehicles and the connected transportation infrastructure will create a new venue for hackers to exploit vulnerabilities to attack. Ethical questions are also involved in programming cars to act in situations in which human injury or death is inevitable, especially when there are split-second choices to be made about whom to put at risk. The legal systems in most states in the US do not have rules covering self-driving cars.

⁵⁶ European Civil Law Rules in Robotics, 2016 at

[http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU\(2016\)571379_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf)

As of 2016, four states in the US (Nevada, Florida, California, and Michigan), Ontario in Canada, the United Kingdom, France, and Switzerland have passed rules for the testing of self-driving cars on public roads. However, these laws do not address issues about responsibility and assignment of blame for an accident for self-driving and semi self-driving cars.

For instance, Nevada passed a law broadly permitting autonomous vehicles and instructed the Nevada Department of Motor Vehicles to craft requirements. Meanwhile, the National Highway Transportation Safety Administration has determined that a self-driving car system, rather than the vehicle occupants, can be considered the “driver” of a vehicle. Some car designs sidestep this issue by staying in autonomous mode only when hands are on the wheel (at least every so often), so that the human driver has ultimate control and responsibility. Still, Tesla’s adoption of this strategy did not prevent the first traffic fatality involving an autonomous car, which occurred in June of 2016. Such incidents are sure to influence public attitudes towards autonomous driving. And as most people’s first experience with embodied agents, autonomous transportation will strongly influence the public’s perception of AI.

Future litigation and jurisprudence regarding autonomous vehicles liability might run into roadblocks when looking at the limited body of case law. In the past, it was difficult to establish is autonomous vehicle cases, where courts have applied different liability theories. It remains to be seen whether the principles of *res ipsa loquitur* will be used by modern courts to conclude that the autonomous vehicle, and not the driver/operator, is at fault. Defendants will argue that the doctrine should not apply when it is unreasonable to infer that the accident was caused by a design or manufacturing defect, or when the accident in question is not one ordinarily seen with design flaws. Complex questions will continue to arise when autonomous vehicles are involved in accidents and/or cause injury.⁵⁷

⁵⁷ Artificial Intelligence Litigation: Can the Law Keep Pace with The Rise of the Machines?, at <http://www.quinnemanuel.com/the-firm/news-events/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/>. See Restatement (Third) of Torts: Prod. Liab. § 3 (1998). Also, in the US case *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, when the court found equal credibility in the defense theory that the loading of the plane was improper, such that a strong gust of wind caused the plane to crash. Even cases decided almost fifty years ago reflect the current legal analysis concerning the question of liability for automated technologies. Also, in the US case *Nelson v. American Airlines, Inc.*, 70 Cal. Rptr. 33 (Cal. Ct. App. 1968), the Court applied the doctrine of *res ipsa loquitur* in finding an inference of negligence by American Airlines relating to injuries suffered while one of its planes was on autopilot, 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. More recently, auto manufacturer Toyota was embroiled in a multi-district litigation matter involving allegations that certain of its vehicles had a software defect that caused the vehicles to accelerate notwithstanding measures the drivers took to stop. The court denied Toyota’s motion for summary judgment premised on the grounds that there could be no liability, because the plaintiff and plaintiff’s experts were unable to identify a

CRIMINAL LIABILITY

As AI is organized to directly interact with the world, one can expect greater invocation of liability for harms caused by AI. The prospect that AI will behave in ways designers did not intend challenges the prevailing assumption within tort law that courts only compensate for foreseeable injuries. *Courts might arbitrarily assign liability to a human actor even when liability is better located elsewhere for reasons of fairness or efficiency. Alternatively, courts could refuse to find liability because the defendant before the court did not, and could not, foresee the harm that the AI caused. Liability would then fall by default on the blameless victim.* The role of product liability, and the responsibility that falls to manufacturers, will likely grow when human actors become less responsible for the actions of a machine.

If tort law expects harm to be foreseeable, criminal law goes further to expect that harm was intended. All legal systems attach great importance to the concept of *mens rea* (mental state). As AI applications engage in behavior that, were it done by human, would constitute a crime, courts and other legal actors will have to puzzle through whom to hold accountable and on what theory, or indeed, whether to hold any individual accountable for consequences that he or she may not have specifically intended, or to have foreseen or authorized.

Box 5. Online bots' criminal behavior

CNBC reported an incident involving online “bots,” where an “automated online shopping bot” was set up by a Swiss art group, given a weekly allowance of \$100 worth of Bitcoin—an online cryptocurrency—and programmed to purchase random items from the “dark web” where shoppers can buy illegal/stolen items. In January 2015, the Swiss police confiscated the robot and its illegal purchases to date, but did not charge the bot or the artists who designed it with any crime.⁵⁸

ASSESS PRIVACY LAW IN LIGHT OF THE BENEFITS OF AI

In an era of increasing data collection and use, privacy protection is more important than ever before. To foster advances in AI that benefit society, policy frameworks must protect privacy without limiting innovation. For example, governments should encourage the development of anonymization techniques that enable analysis of large data sets without revealing individual identities and enact laws that recognize the value of anonymization in preserving privacy. Privacy laws should also recognize that data collected for a particular purpose may lead to beneficial AI insights. To support useful research, governments should provide reasonable latitude in assessing whether data used for AI analysis is within the scope of its original purpose. While privacy law

precise software design or manufacturing defect, instead finding that the evidence supported inferences from which a reasonable jury could conclude that the vehicle continued to accelerate and failed to slow or stop despite the plaintiff's application of the brakes. In re Toyota Motor Corp. Unintended Acceleration Mktg., Sales Practices, & Prod. Liab. Litig., 978 F. Supp. 2d 1053, 1100-01 (C.D. Cal. 2013).

⁵⁸ <https://www.cnn.com/2015/04/21/robot-with-100-bitcoin-buys-drugs-gets-arrested.html>

should account for the benefits of AI, new regulations may be required to address concerns about the predictive power of AI, such as the possibility that AI systems may infer private information about people.

VI. THE ETHICS OF AI

ENCOURAGE DEVELOPMENT OF ETHICAL BEST PRACTICES

AI may inherently predict data or authorize actions based on its “understanding” of patterns in underlying data, and the values embodied in its learning and thinking approach. Questions have already arisen concerning the use of AI and the law. In many cases, the law does not yet provide guidance. In some cases, the legislation may not be able to address AI functionality that overlaps areas that historically have been reserved to areas of personal morality and belief. Hence, it is generally agreed that ethical implications will play an important role in governing AI activities, and in establishing the standard of care expected of practitioners and clients in these areas.

The Black Box Problem

Because AI systems will assume responsibility from humans – and for humans – it is important that people understand how these systems might fail. However, this does not always happen in practice.⁵⁹

Consider the Northpointe algorithm that US courts used to predict reoffending criminals. The algorithm weighed 100 factors such as prior arrests, family life, drug use, age and sex, and predicted the likelihood that a defendant would commit another crime. Northpointe’s developers did not specifically consider race, but when investigative journalists from ProPublica analyzed Northpointe, it found that the algorithm incorrectly labeled black defendants as “high risks” almost twice as often as white defendants. Unaware of this bias and eager to improve their criminal justice system, states like Wisconsin, Florida, and New York trusted the algorithm for years to determine sentences. Without understanding the tools they were using, these courts incarcerated defendants based on flawed calculations.⁶⁰

The Northpointe case offers a preview of the potential dangers of deploying AI systems that people do not fully understand. Current machine-learning systems operate so quickly that no one really knows how they make decisions – not even the people who develop them. Moreover, these systems learn from their environment and update their behavior, making it more difficult for researchers to control and understand the decision-making process. This lack of transparency –

⁵⁹ Towards a Code of Ethics in Artificial Intelligence with Paula Boddington, Davey, Future of Life Institute, at <https://futureoflife.org/2017/07/31/towards-a-code-of-ethics-in-artificial-intelligence/>.

⁶⁰ Ibid.

the “black box” problem – makes it extremely difficult to construct and enforce a code of ethics (see also Box 3 for use of algorithms in judicial decisions).

Box 6. The use of algorithms to make judicial decisions has provided a host of problematic examples

In Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington and Wisconsin, judges can use such algorithms in determining criminal sentencing. The Supreme Court in Wisconsin ruled recently in the 2016 case of *State of Wisconsin v Eric Loomis* that the use of the COMPAS algorithm used to help determine sentence length was constitutional.

The European Parliament is currently considering issues of law and robotics in its report on European Civil Law and Robotics issued in January 2017. The report similarly notes the range of ethical issues and judgments that AI may implicate. For example, AI could be used to invade personal privacy by accurately inferring information that people would prefer to keep private. It can inadvertently perpetuate discrimination. AI may also be required to make difficult moral choices such as when driverless cars have to pick whom to injure when seeking to avoid an accident.

These and other cases suggest that transparency about AI analysis will be critical in any ethical framework. Governments, industry, and civil society should begin to work together to weigh the range of ethical issues that AI raises, with the goal of developing guidelines and best practices. As experience with AI broadens, it may make sense to establish more formal industry standards that reflect consensus about ethical issues but that do not impede innovation and progress in the development of AI capabilities.

A number of important ethical issues may be implicated.

PROPOSED ANALYTICAL FRAMEWORK

Should crimes and misdemeanor be assessed through AI? It would be more systematic but would be a sophisticated form of mandatory sentences, with all the drawback of that kind of systems. Should AI devices be able to inflict punishments on human beings? It is possible to set better boundaries for machines.

(i) Ethical programming: Can AI devices be programmed (and trusted) to make ethical decisions?

For instance, the Milgram experiment where people were induced to wish to send lethal doses of electric shocks to other people to make them learn and correct their mistakes, would not happen with an AI device, if ethical programming is included in the AI programming. The question is what happens if the ethical programming is bug ridden. To some extent, the Milgram experiment shows “ethical bugging” by human beings.

(ii) Moral decisions that were made on the spot, in particular in emergency situations, now have to be pre-programmed and planned.

One example is the “tunnel problem” which is a variation on the “trolley problem” introduced in the 1970s by Philippa Foot. The “tunnel problem” wonders how an automatic car is supposed to react when it has the choice between killing its passenger and killing a pedestrian who is unexpectedly crossing the road in front of the car.

(iii) Infringement of privacy.

An AI device has no inherent notion of privacy or general principles of human dignity yet, if we develop the notion of ethical programming, it may be possible to do so (e.g. a surveillance camera could switch off when people are undressing).

An important question is whether policymakers should, or can, legislate these issues directly, or can only insist on transparency regarding the ethical principles adopted.

(iv) Risk of loss of control over AI devices.

With AI devices, humans are losing the ability to control the machine.

When AI devices start producing AI devices, there is possibility for losing human control over the creation of AI devices. Self-replication of some AI devices creates the risk of chain reactions (e.g. some internet viruses before the development of more powerful antiviruses), or mutually reinforcing actions and responses.

A new generation of AI devices is trying to clone people’s voices to record things they did not say. This development suggests that, in the near future, AI devices may effectively ascribe fault or behavior for their own actions to individuals. Rules for evidence, discovery and judicial procedure may be implicated.

(v) The danger of embedded and amplified discriminations.

Many AI devices are better than human beings at identifying small differences. This can be used to our advantage (e.g. apps differentiating cancerous skin growth from other skin spots or moles). However, “black box” algorithms and machine learning may also develop (or embody) false correlations between appearance, origin or other human attributes, that replicate and extend discriminatory practices.

Box 7. The problem of algorithmic bias

(i) Microsoft Tay was designed to engage with people ages 18 to 24, and it burst onto social media with an upbeat "helllllooooo world!!" (the "o" in "world" was a planet earth emoji). But within 12 hours, Tay morphed into a foul-mouthed racist Holocaust denier that said feminists "should all die and burn in hell." and was quickly removed from Twitter. Tay was programmed to learn from the behaviors of other Twitter users, and in that regard, the bot was a success. Tay's embrace of humanity’s worst attributes is an example of algorithmic bias—when seemingly innocuous programming takes on the prejudices either of its creators or the data it is fed.

(ii) In 2015, Google Photos tagged several African-American users as gorillas, and the images lit up social media. Yonatan Zunger, Google's chief social architect and head of

infrastructure for Google Assistant, quickly took to Twitter to announce that Google was scrambling a team to address the issue.

One of the trickiest parts about algorithmic bias is that engineers don't have to be actively racist or sexist to create it. In an era when we increasingly trust technology to be more neutral than we are, this is a dangerous situation. As the tech industry begins to create artificial intelligence, it risks inserting racism and other prejudices into code that will make decisions for years to come. And as deep learning means that code, not humans, will write code, there is an even greater need to root out algorithmic bias. Put simply, the best case against AI is that it learns from us (humans) and we are awful! Thus, the ongoing diversity and inclusion efforts in this area are of great importance⁶¹.

VII. THE WAY FORWARD

AI raises a series of complex issues that cut across social, economic, political, technological, legal, ethical and philosophical boundaries. To untangle the uncertainties, possibilities and potential perils, there is a need to assess and understand the correlation between these fields. On one hand, AI offers enormous promise of unleashing wealth creation. On the other hand, AI may also tend to exacerbate income inequality and the “digital divide” within nations and between nations. Policies and regulatory frameworks even in developed countries have not been able to keep pace with technological advances.

The likely outcome, at least in the interim, may be a legal framework largely comprising codes of conduct, voluntary self-regulation, and evolving contractual practices. Cutting edge businesses in particular are expected to rely on contracts to govern various aspects of their business in absence of existing laws. Lawyers will inevitably play a key role in developing these underpinnings of a sound, secure and appropriate AI economy that fully realizes this new source of innovation.

Accordingly, it is imperative for the World Bank Group to have a strong group of thinking, well-trained practitioners on these issues, who can counsel clients and inform public debate.

It is equally important then for World Bank Group practitioners to maintain a mechanism that keeps them on the cutting edge of policy thinking and legal developments in this area.

Key points to consider:

- AI should be used for the public good

⁶¹ <https://www.wired.com/2017/02/keep-ai-turning-racist-monster/>

- Governments should embrace AI by developing a human-centered governance approach to Industry 4.0
- If and how AI should be regulated (The US White House generated a report of ‘Preparing for the Future of Artificial Intelligence,’ and a companion “National Artificial Intelligence Research and Development Strategic Plan,” in 2016. The White House also co-hosted public workshops on AI policy areas and requested information from the public on AI issues. Japan has pushed for basic rules on AI at the G7 meetings in 2016).
- Ethics framework for AI should be developed (South Korea is developing a robot ethics charter)
- The AI should be used to supplement human workers, not replace them
- Developing countries should participate in take advantage of AI advances, through implementation of human centered governance model in Industry 4.0
- Data should be free from bias⁶²

World Bank’s AI agenda should be driven by the following considerations

Foster generation and proliferation of upstream knowledge in AI. In accordance with WBG’s mission of shared prosperity, WBG should support mechanisms that unlock the potential of upstream knowledge to create solutions for broader economic development. WBG’s role in initiatives like GAVI etc. could serve as a template for interventions in AI. Specifically, business models built on upstream knowledge should be supported for greater proliferation of technological solutions reliant on AI.

WBG should use its convening power to advocate for use of AI for economic development and play a key role in shaping thinking on best practices on adopting AI for poverty alleviation and shared prosperity

WBG should develop finance and advisory products in addition to knowledge stock that support greater access to AI by developing countries

WBG should use its vast global experience to promote appropriate regulation/self-regulation mechanisms that strike the appropriate balance between mitigating risks posed by AI and the enormous potential for greater well-being.

⁶² Adapted from <https://futurism.com/images/white-house-ai-report-everything-you-need-to-know-infographic/>

ANNEX 1. AI CASE LAW: USA

Agency law

As AI programs become more adaptive and capable of learning on their own, courts will have to determine whether such programs can be subject to a unique variant of agency law. Current laws of agency may not apply, because 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) (“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.”); 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). As a result, it is possible that the courts or legislatures will be asked to impose strict liability on the creators of programs, for the acts of such programs.

Product liability

Product liability claims and conventional views of culpability and ethics are certain to be tested by these autonomous machines—like self-driving vehicles—where the current roadmap is for a mixed human and AI driver world. Product liability law provides some framework for resolving such claims; 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. Complications may arise when product liability claims are directed to failures in software, as computer code has not generally been considered a “product” but instead is thought of as a “service,” with cases seeking compensation caused by alleged defective software more often proceeding as breach of warranty cases 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).

Under these metrics, courts will have to assess what liability to impose for accidents involving the various types of automated vehicles available today, as well as those soon to be released. **One option is to insist on strict liability for manufacturers of the automated systems. If there is no strict liability, a court might find itself in uncharted waters if forced to make a determination as to how best to weigh the comparative liability of AI programs and drivers.** The solution suggested by the existing law, while dated, would hold the vehicle’s manufacturer liable and let the manufacturer seek indemnity or contribution from other parties, if any, that might be responsible. However, consideration also may be given to apportioning responsibility among all of the parties that participated in building and maintaining the vehicle’s autonomous systems, through the application of a variation of “common enterprise” liability. 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 (9th Cir. 2010); *SEC v. R.G. Reynolds Enters., Inc.*, 952 F.2d 1125 (9th Cir. 1991); *FTC v. Tax Club, Inc.*, 994 F. Supp. 2d 461 (S.D.N.Y. 2014). **A “common enterprise” theory might allow the law to impose joint liability, for limited types of claims, without having to assign every aspect of wrongdoing to one party or another. Another issue to consider is the issue of corporate governance and Board responsibility. At least until millennials grow into leadership there will be a concern for Board competence and reliance on third party experts, and whether Board will be held accountable for what it does not really understand.**

Legislatures and regulatory agencies have already been making great strides to determine how best to attribute fault in such situations. For example, the states of Nevada, Florida, California, Michigan and Tennessee and the District of Columbia have all passed legislation related to autonomous automobiles, and nineteen additional states have similar bills under consideration. See Jessica S. Brodsky, *Autonomous Vehicle Regulation: How an Uncertain Legal Landscape May Hit the Brakes on Self-Driving Cars*, 31 *Berkeley Tech. L.J.* 851 (2016). Sophisticated parties are destined to address a variety of complicated legal issues presented with the advent of AI technologies and products. In particular, the competing interests between manufacturers of various AI components and the end products that incorporate those components will need to be addressed through contracts and robust indemnification agreements. Legislators and courts will soon have to answer the questions such as whether a machine can enter into a binding contract on behalf of itself, or a person it represents, and does a machine-negotiated contract redefine what it means to look to the understanding of one party or between parties? We are at the precipice of requiring new definitions for scienter, “meeting of the minds,” and a host of other black letter law constructs that have served as the underpinning of commercial litigation for generations.

Replacing Professional Judgment with Computers: Malpractice Claims Anticipated

There is no dispute that the legal and medical professions are among the professions that require the greatest decision-making and exercise of judgment. It is because of this that claims of malpractice are available to those who rely on the decision-making and judgment of the skilled, trained professionals who practice in these fields. It is also the case that these are two fields that are introducing an increasing number of AI-based technologies. In the legal industry, a growing interest in “big data” and natural language processing has resulted in start-ups seeking to tackle the difficult task of aggregating, synthesizing and modeling a collective corpus of case law. One example, RavelLaw uses natural language processing to identify, extract and classify information from legal documents, automating basic case law analysis to make research more efficient and targeted. The company hopes to add automated analysis of briefs, wording recommendations for particular judges, and probability-based outcome predictions to litigators and their clients. Another, ROSS Intelligence calls itself “Your Brand New Artificially Intelligent Lawyer” and is built in partnership with IBM using the Watson artificial intelligence supercomputer. The company highlights its ability to process natural language to assist in case law review. Another area that has had significant penetration within law firms and with clients, is the use of AI to

review documents. The advent of e-discovery is such that it is not as efficient, or economical, to have attorneys conduct first reviews of the massive volumes of documents collected in large litigations. Attorney oversight remains necessary, in particular to guarantee adequate controls are in place to secure privileged and confidential information from inadvertent disclosures.

In the medical industry, robotic surgical instruments and cancer treatment devices, as well as the continued development and adoption of IBM's Watson for medical treatment has led to increased analysis of potential liability for the use of such instruments and devices. As mentioned above, there is precedent for litigation over the safety of surgical robots, with the claims all proceeding on some form of agency theory, rather than claiming that the robot itself bears liability. By combining elements from medical malpractice, vicarious liability, products liability, and enterprise liability, the law can create a uniform approach for AI systems, thereby eliminating any inequities that may arise from courts applying different theories of liability and encouraging the continued beneficial use of such systems.

Medical malpractice is applied to healthcare providers, while vicarious liability tends to focus on institutions that employ healthcare providers. It is possible to envision a medical malpractice action based on a lack of informed consent arising when a physician fails to inform the patient of all relevant information about a course of treatment, including any risks associated with the use of autonomous machines for such treatment. The hospital's own duty to supervise the quality of medical care administered in the facility would be related to actions asserting vicarious liability, so long as the court determines that the autonomous machine can be analogized to an employee. If a court decides instead to analogize the AI system to a machine like a Magnetic Resonance Imaging device, then products liability claims may be attached to defective equipment and medical devices that healthcare providers may use. While manufacturers of medical equipment and devices can be liable through products liability actions, the learned intermediary doctrine results in the manufacturer having no duty to the patient and thus prevents plaintiffs from suing medical device manufacturers directly. See, e.g. *Banker v. Hoehn*, 278 A.D.2d 720, 721, 718 N.Y.S.2d 438, 440 (2000). This liability structure makes it challenging for patients to win products liability suits in medical device cases.

While AI innovations are certain to save time and money, there are concerns that AI technology, when used to replace human professional judgment, could lead to increased claims raising complex issues of causation, legal duties, and also liability. A regime based on some form of enterprise liability, similar to what has been discussed previously in relation to autonomous vehicles, which combines elements of malpractice, products liability, and vicarious liability, could address these legal challenges while encouraging professionals to purchase and use these AI systems.⁶³

⁶³ <http://www.quinnemanuel.com/the-firm/news-events/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/>

ANNEX 2 - BIBLIOGRAPHY

ARTIFICIAL INTELLIGENCE AND LIFE IN 2030 ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016 @ https://ai100.stanford.edu/sites/default/files/ai100report10032016fnl_singles.pdf

Policy Recommendations, Artificial Intelligence @ https://news.microsoft.com/cloudforgood/_media/downloads/en/artificial-intelligence-en.pdf

Is Artificial Intelligence Protectable by Law? Should it Be? @ <http://www.thefashionlaw.com/home/is-artificial-intelligence-protectable-by-law-should-it-be>

When AI Goes Bad: Case Studies @ <https://www.cs.ox.ac.uk/efai/towards-a-code-of-ethics-for-artificial-intelligence/what-are-the-issues/ai-ethics-fails-case-studies/>

European Civil Law Rules in Robotics, 2016 @ [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU\(2016\)571379_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf)

DRAFT REPORT with recommendations to the Commission on Civil Law Rules on Robotics, European Parliament, 2016 @ <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2f%2fEP%2f%2fNONSGML%2bCOMPARL%2bPE-582.443%2b01%2bDOC%2bPDF%2bV0%2f%2fEN>

HOW TO KEEP YOUR AI FROM TURNING INTO A RACIST MONSTER @ <https://www.wired.com/2017/02/keep-ai-turning-racist-monster/>

Artificial Intelligence Litigation: Can the Law Keep Pace with The Rise of the Machines? @ <http://www.quinnemanuel.com/the-firm/news-events/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/>

Robot with \$100 bitcoin buys drugs, gets arrested @ <http://www.cnn.com/2015/04/21/robot-with-100-bitcoin-buys-drugs-gets-arrested.html>

Give robots 'personhood' status, EU committee argues @ <https://www.theguardian.com/technology/2017/jan/12/give-robots-personhood-status-eu-committee-argues>

Coming to Grips with Artificial Intelligence, Law Library Lights, Vol. 60, Issue 2 (Winter 2017), pp. 9-10, Zimmerman, Matt

Artificial Intelligence: Not Just Sci-Fi Anymore, AALL Spectrum, Vol. 21, Issue 1 (September/October 2016), pp. 34-37 Gediman, Mark

Peeping HALs: Making Sense of Artificial Intelligence and Privacy, European Journal of Legal Studies, Vol. 2, Issue 3 (2010), pp. 168-192 Calo, Ryan

Patents in An Era of Infinite Monkeys and Artificial Intelligence, Stanford Technology Law Review, Vol. 19, Issue 1 (Fall 2015), pp. 32-51

Carpooling Liability: Applying Tort Law Principles to the Joint Emergence of Self-Driving Automobiles and Transportation Network Companies / Walpert, Jacob D., *Fordham Law Review*, Vol. 85, Issue 4 (March 2017), pp. 1863-[ii]

Legal Technology: Artificial Intelligence and the Future of Law Practice / McKamey, Mark, *Appeal: Review of Current Law and Law Reform*, Vol. 22 (2017), pp. 45-58

Artificial Intelligence in the Practice of Law: An Analysis and Proof of Concept Experiment / Ben-Ari, Daniel; Frish, Yael; Lazovski, Adam; Eldan, Uriel; Greenbaum, Dov, *Richmond Journal of Law & Technology*, Vol. 23, Issue 2 (2017), pp. 2-55

Artificial Intelligence and the Copyright Dilemma / Hristov, Kalin, *IDEA: The Journal of the Franklin Pierce Center for Intellectual Property*, Vol. 57, Issue 3 (2017), pp. 431-454

Artificial Legal Intelligence on the Internet: The Next Approach to Enforcing the Law Online / Zatarain, Jesus Manuel Niebla, *Edinburgh Student Law Review*, Vol. 2, Issue 4 (2015), pp. 64-75

Siri-Ously? Free Speech Rights and Artificial Intelligence / Massaro, Toni M.; Norton, Helen, *Northwestern University Law Review*, Vol. 110, Issue 5 (2016), pp. 1169-1194

Machines without Principals: Liability Rules and Artificial Intelligence / Vladeck, David C., *Washington Law Review*, Vol. 89, Issue 1 (March 2014), pp. 117-150

Artificial Intelligence: Robots, Avatars, and the Demise of the Human Mediator / Larson, David Allen, *Ohio State Journal on Dispute Resolution*, Vol. 25, Issue 1 (2010), pp. 105-164

Editorial: The Future of IP Law in an Age of Artificial Intelligence / Schafer, Burkhard; *SCRIPTed: A Journal of Law, Technology and Society*, Vol. 13, Issue 3 (December 2016), pp. 283-288

Computers as Inventors - Legal and Policy Implications of Artificial Intelligence on Patent Law / Fraser, Erica, *SCRIPTed: A Journal of Law, Technology and Society*, Vol. 13, Issue 3 (December 2016), pp. 305-333

Expert Robot: Using Artificial Intelligence to Assist Judges in Admitting Scientific Expert Testimony / Katz, Pamela S., *Albany Law Journal of Science & Technology*, Vol. 24, Issue 1 (2014), pp. 1-46

The Criminal Liability of Artificial Intelligence Entities - From Science Fiction to Legal Social Control / Hallevy, Gabriel, *Akron Intellectual Property Journal*, Vol. 4, Issue 2 (2010), pp. 171-202

From Jeopardy to Jaundice: The Medical Liability Implications of Dr. Watson and Other Artificial Intelligence Systems / Allain, Jessica S., *Louisiana Law Review*, Vol. 73, Issue 4 (Summer 2013), pp. 1049-1080

Four Futures of Legal Automation / Pasquale, Frank; Cashwell, Glyn, *UCLA Law Review Discourse*, Vol. 63 (2015), pp. 26-48 63

Trial by Machine / Roth, Andrea, Georgetown Law Journal, Vol. 104, Issue 5 (June 2016), pp. 1245-1306

The New Investor / Lin, Tom C. W., UCLA Law Review, Vol. 60, Issue 3 (February 2013), pp. 678-735

Minimizing Global Catastrophic and Existential Risks from Emerging Technologies through International Law / Wilson, Grant, Virginia Environmental Law Journal, Vol. 31, Issue 2 (2013), pp. 307-364

Robotics and the Lessons of Cyberlaw / Calo, Ryan, California Law Review, Vol. 103, Issue 3 (June 2015), pp. 513-564

Preparing Financial Regulation for the Second Machine Age: The Need for Oversight of Digital Intermediaries in the Futures Markets / Scopino, Gregory, Columbia Business Law Review, Vol. 2015, Issue 2 (2015), pp. 439-519

From Cyberspace to outer Space: Existing Legal Regimes under Pressure from Emerging Meta-Technologies / Lauria, Rita M.; Robinson, George S., University of La Verne Law Review, Vol. 33, Issue 2 (May 2012), pp. 219-254

Algorithmic Choice and Superior Responsibility: Closing the Gap between Liability and Lethal Autonomy by Defining the Line between Actors and Tools / Reitingner, Nathan, Gonzaga Law Review, Vol. 51, Issue 1 (2015/2016), pp. 79-120

Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies / Matthew U. Scherer , 29 Harv. J. Law & Tec 353 (Spring, 2016)

Artificial Intelligence and the Law: Essay: The Scored Society: Due Process for Automated Predictions / Danielle Keats Citron and Frank Pasquale , 89 Wash. L. Rev. 1 (March, 2014)

Pulse Symposium: Imagining the Legal Landscape: Technology and the Law in 2030: Policing Police Robots / Elizabeth E. Joh, 64 UCLA L. Rev. Disc. 516 (2016)

"I, Robot - I, Criminal"--When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offenses / Gabriel Hallevy, 2010 Syracuse Sci. & Tech. L. Rep. 1 (Spring 2010)

Analysis of the Legal Status of Unmanned Commercial Vessels in U.S. Admiralty and Maritime Law / Michal. Chwedczuk 47 J. Mar. L. & Com. 123 (April, 2016)

AI Start-up 'Legit' to Disrupt Patent World, Secures Seedcamp Funding 25th November 2016 artificiallawyer @ <https://www.artificiallawyer.com/2016/11/25/ai-start-up-legit-to-disrupt-patent-world-secures-seedcamp-funding/>

Europe is leading the way in AI and machine learning (and even Silicon Valley wants in) @ http://www.wired.co.uk/article/deep-tech-europe-hubs?utm_content=bufferd3d59&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

The Global Race for Artificial Intelligence – Comparison of Patenting Trends @

<https://www.wilsoncenter.org/blog-post/the-global-race-for-artificial-intelligence-comparison-patenting-trends>

IP for AI: can we patent an artificial human expert? @

<http://www.lexology.com/library/detail.aspx?g=b150f084-2a4c-4959-8505-8fd7ddcb8e9d>

Awe and anxiety: what artificial intelligence means for IP @

<http://www.worldipreview.com/article/awe-and-anxiety-what-artificial-intelligence-means-for-ip>

Machine learning versus AI: what's the difference? @

<http://www.wired.co.uk/article/machine-learning-ai-explained>

<http://robohub.org/should-an-artificial-intelligence-be-allowed-to-get-a-patent/>

ANNEX 3 - WHO IS WHO

Association for the Advancement of Artificial Intelligence (AAAI) <http://www.aaai.org/>

World Economic Forum's Council on the Future of Artificial Intelligence and Robotics

Jose Angelo Junior Gurzoni, a Senior IT Officer at the IFC,

http://isearch.worldbank.org/skillfinder/ppl_profile_new/000229216

Kerstin Vignard, Chief of Operations and Deputy to the Director for the United Nations Institute for Disarmament Research, <http://www.unidir.org/about/staff/kerstin-vignard>.

Francis Gurry, Director General, World Intellectual Property Office

Choi Dong-gyou, Commissioner, Korea Intellectual Property Office

Meredith Whittaker, Founder and Lead, Google Open Research

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/meredith-whittaker>

LinkedIn: <https://www.linkedin.com/in/meredith-whittaker-4a57b94>

Kate Crawford, Principal Researcher / Senior Research Fellow, Microsoft Research / NYU

@katecrawford

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/kate-crawford>

LinkedIn: <https://www.linkedin.com/in/kate-crawford-a66574>

Website: <https://www.microsoft.com/en-us/research/people/kate/>

Kay Firth-Butterfield, Chief Officer, Ethics Advisory Panel, Lucid AI

@KayFButterfield

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/kay-firth-butterfield>

LinkedIn: <https://www.linkedin.com/in/kay-firth-butterfield-b5b3853a>

Tim Hwang, Public Policy Counsel, Google

@timhwang

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/tim-hwang>

LinkedIn: <https://www.linkedin.com/in/timhwang>

Yann LeCunn, Director of AI research, Facebook

@ylecun

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/yann-lecunn>

LinkedIn: <https://www.linkedin.com/in/yann-lecun-0b999>

Website: <http://yann.lecun.com/>

Carolyn Nguyen, Director, Technology Policy, Microsoft

@mhcnguyen

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/carolyn-nguyen>

LinkedIn: <https://www.linkedin.com/in/mhcarolynn>

Chris DiBona, Director, Open Source & Science Outreach, Google

@cdibona

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/chris-dibona>

LinkedIn: <https://www.linkedin.com/in/cdibona>

Christian Sandvig, Professor, University of Michigan

@niftyc

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/christian-sandvig>

LinkedIn: <https://www.linkedin.com/in/niftyc>

Claudia Perlich, Chief Scientist, Dstillery/NYU

@claudia_perlich

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/claudia-perlich>

LinkedIn: <https://www.linkedin.com/in/cperlich>

Cynthia Braezeal, Founder + Chief Scientist/ Assoc. Professor Media Arts and Sciences

Jibo/ MIT Media Lab

@cynthiabraezeal

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/cynthia-braezeal>

LinkedIn: <https://www.linkedin.com/in/cynthia-braezeal-1792317>

Danielle Citron, Professor of Law, University of Maryland Carey School of Law

@daniellecitron

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/danielle-citron>

Dawn Nafus, Senior Research Scientist, Intel

@dawnnafus

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/dawn-nafus>

Eric Horvitz, Technical Fellow & Managing Director, Microsoft Research

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/eric-horvitz>

LinkedIn: <https://www.linkedin.com/in/erichorvitz>

Erik Brynjolfsson, Director, MIT Initiative on the Digital Economy; Schussel Family Professor, MIT-Sloan, MIT

@erikbryn

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/erik-brynjolfsson>

Website: <http://ebusiness.mit.edu/erik/>

Fernando Diaz, Senior Researcher, Microsoft Research

@fdiaz_msr

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/fernando-diaz>

LinkedIn: <https://www.linkedin.com/in/fernando-diaz-b6b9023>

Francesca Rossi, Research Scientist and University Professor, IBM T. J. Watson Research Center and University of Padova

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/francesca-rossi>

Interview: <https://futureoflife.org/2017/01/26/francesca-rossi-interview/> from Beneficial AI 2017 Conference

Frank Pasquale, Professor of Law, University of Maryland

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/frank-pasquale>

Website: <https://www.law.umaryland.edu/faculty/profiles/faculty.html?facultynum=984>

LinkedIn: <https://www.linkedin.com/in/frankpasquale>

Genevieve Bell, Sr. Fellow, VP Corporate Strategy, Intel Corp.

@feraldata

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/genevieve-bell>

LinkedIn: <https://au.linkedin.com/in/genevievebell>

Guru Banavar, Vice President, Cognitive Computing, IBM Research, IBM

@banavar

Full bio: <https://artificialintelligencenow.com/schedule/workshop/attendee/guru-banavar>

LinkedIn: <https://www.linkedin.com/in/banavar>

Hanna Wallach, Senior Researcher, Microsoft Research

@hannawallach

Website: <https://www.microsoft.com/en-us/research/people/wallach/>

Heather Roff, Senior Research Fellow at Department of Politics and International Relations
University of Oxford

@hmroff

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/heather-roff>

LinkedIn: <https://uk.linkedin.com/in/heather-m-roff-95466713>

Helen Nissenbaum, Director, New York University Information Law Institute

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/helen-nissenbaum>

Ian Kerr, Canada Research Chair in Ethics, Law and Technology, University of Ottawa

@ianrkerr

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/ian-kerr>

Website: <https://arts.uottawa.ca/philosophy/people/kerr-ian-r>

Jennifer Chayes, Distinguished Scientist, Managing Director of Microsoft Research New
England and Microsoft Research New York City, Microsoft

@jenniferchayes

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/jennifer-chayes>

Website: <https://www.microsoft.com/en-us/research/people/jchayes/>

LinkedIn: <https://www.linkedin.com/in/jennifer-chayes-6328145>

Lucy Suchman, Professor, Anthropology of Science and Technology, Lancaster University, UK

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/lucy-suchman>

Website: <http://www.lancaster.ac.uk/sociology/about-us/people/lucy-suchman>

Madeleine Clare Elish, Researcher, Data & Society

@mcette

LinkedIn: <https://www.linkedin.com/in/mcelish>

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/madeleine-clare-elish>

Manuela Veloso, Herbert A. Simon University Professor, Carnegie Mellon University

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/manuela-veloso>

Website: <http://www.cs.cmu.edu/~mmv/>

Mustafa Suleyman, Co-founder and Head of Applied Artificial Intelligence, Google DeepMind

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/mustafa-suleyman>

LinkedIn: <https://uk.linkedin.com/in/mustafa-suleyman-54543a72>

Paul Dourish, Chancellor's Professor of Informatics, UC Irvine

@dourish

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/paul-dourish>
Websites: <http://www.dourish.com/> and <http://www.ics.uci.edu/~jpd/students.shtml>

Peter Asaro, Assistant Professor, New School, Stanford University
@PeterAsaro

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/peter-asaro>
LinkedIn: <https://www.linkedin.com/in/peterasaro>

Ronald C. Arkin, Regents' Professor and Associate Dean of Research, Georgia Institute of Technology
@RonaldArkin

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/ron-arkin>
Website: <http://www.cc.gatech.edu/aimosaic/faculty/arkin/>

Ryan Calo, Co-founder, University of Washington Tech Policy Lab

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/ryan-calo>
Website: <https://www.law.washington.edu/directory/profile.aspx?ID=713>

Sean Legassick, Policy Advisor, Google DeepMind
@seanlegassick

LinkedIn: <https://uk.linkedin.com/in/seanlegassick>

Seth Goldstein, Assoc. Professor, Carnegie Mellon University
Website: <https://www.cs.cmu.edu/directory/sethg>

Solon Barocas, Postdoc Researcher, Microsoft Research
@s010n

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/solon-barocas>
Website: <http://solon.barocas.org/>

Sue Glueck, Academic Relations Director, Microsoft
@sueglueck

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/sue-glueck>
LinkedIn: <https://www.linkedin.com/in/sue-glueck-33746510>

Suresh Venkatasubramanian, Associate Professor, University of Utah
@geomblog

Full bio: <https://artificialintelligenow.com/schedule/workshop/attendee/suresh-venkatasubramanian>
LinkedIn: <https://www.linkedin.com/in/suresh-venkatasubramanian-233b751>
Website: <http://www.cs.utah.edu/~suresh/web/>

GLOBAL FORUM on LAW, JUSTICE and DEVELOPMENT

c/o The World Bank

1818 H Street, NW

Washington, DC 20433

Mail stop MC6-607

Email: globalforumljd@worldbank.org

Web: www.globalforumljd.org

