Masthead Logo

Santa Clara High Technology Law Journal

Volume 35 | Issue 4 Article 1

4-21-2019

ARTIFICIAL INTELLIGENCE AND THE FOG OF INNOVATION: A DEEP-DIVE ON GOVERNANCE AND THE LIABILITY OF AUTONOMOUS SYSTEMS

Brandon W. Jackson

ARTIFICIAL INTELLIGENCE AND THE FOG OF INNOVATION: A DEEP-DIVE ON GOVERNANCE AND THE LIABILITY OF AUTONOMOUS SYSTEMS

By Brandon W. Jackson [†]

ABSTRACT

The convergence of various technological advancements has caused numerous industries to pivot towards the development of artificial intelligence and machine learning. With promises to positively augment nearly all aspects of our daily lives, artificial intelligence and machine learning have the potential to change the very fabric of our society. To realize these benefits, or perhaps avoid the forecasted perils of artificial intelligence, society must overcome certain technological, legal, and social challenges. The complexity of these issues is compounded by the fact that AI does not exist in a vacuum, and it is intertwined with data rights, privacy rights, intellectual property protection, ethics, civil rights, due process, geopolitics, and social values, among other considerations. As with many areas of innovation, the technology rapidly outpaces the law.

This paper seeks to further the discussion of governing artificial intelligence systems and the role of liability as an indirect form of regulation. It is premised on the likelihood that broad regulation of the underlying technologies is unlikely in the United States given the relevant infant state of artificial intelligence and the potential social and economic benefits offered by machine learning. This paper supports the contention that judicial decision-making will play a significant role in shaping the artificial intelligence landscape. It examines principles of agency and alternative legal identities as applied to artificial intelligence. Further, this paper explores problems associated with trying to apply common law principles of liability, such as product liability and strict liability, to autonomous systems. As these systems move towards greater autonomy, common law principles of

† The views expressed in this paper are expressly those of the author and do not reflect those of the U.S. Government or Department of Defense. Brandon Jackson is an employee of the Department of Defense and is a Professorial Lecturer in Law at The George Washington University Law School. Special thanks to Professor Paul Rosenzweig for his guidance on previous drafts and Associate Dean Lisa Schenck for her encouragement during the drafting process. Also special thanks to my wife and family for their continued support, as well as my grandmother, Marjorie Lampe, for her editorial contributions to previous drafts.

agency will become too attenuated and assigning liability will prove to be difficult under current liability rules. Thus, the courts will likely struggle to harmonize the nuances of innovation with traditional concepts of law. This paper argues that to capitalize on the benefits offered by artificial intelligence, or more importantly to avoid its destructive perils, it will be paramount for society and the law to evolve.

CONTENTS

INTRODUCTION

Alan Turing, in his famous 1950 paper, "Computing Machinery and Intelligence," wrote, "we can only see a short distance ahead, but we can see plenty there that needs to be done."[1] This sentiment, expressed nearly 70 years ago in the context of whether machines can think, reflects the current momentum of recent technological breakthroughs to endow machines with the ability to make intelligent decisions — the concept of Artificial Intelligence (AI). While the notion of AI is not novel, it has recently become a driving factor in industry because of compounded advancements in the availability of big data, machine learning approaches and algorithms, and powerful computing mechanisms.[2] More importantly, these technological breakthroughs have provided tangible realizations of how AI can be infused into nearly all domains to address society's greatest challenges. Even with these advancements, however, the exploration into AI is seated in infancy as society seeks to understand and overcome the technological, social, and legal challenges of computer systems endowed with human characteristics and abilities.

Turing's sentiments towards progress in AI are not unique to technological development; rather, they stand as a modern summation of the legal and social thinking that continues to be necessitated by society's reach for a scientific way to augment the human experience. This paper is designed to further the discussion of AI governance and, specifically, the role of liability as an indirect form of regulation. Part I examines the technological foundation for AI, as well as the promises and perils it holds, as a precursor to understanding the encompassing issues of law and policy. Part II explores the technological, legal, and social barriers to AI governance, including how governing issues are compounded by the blended nature of AI with other technological domains, such as privacy, big data, and cybersecurity. In light of these challenges, it is likely that judicial decisions surrounding tort liability will be a driving force in shaping the AI landscape. Lastly, Part III analyzes the competencies of traditional liability regimes to remedy harms caused by AI systems. To an extent, the concept of strict liability is the most amenable tort regime that can be harmonized with emerging

---

[1] A.M. Turing, *Computing Machinery and Intelligence* (1950), https://www.csee.umbc.edu/courses/471/papers/turing.pdf.

[2] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 6 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf. (discussing big data, improved machine learning approaches and algorithms, and more powerful computers as three factors that began driving progress and enthusiasm for AI around 2010).

AI technologies. However, as the technology pushes towards greater autonomy in effectuating action, legal principles of agency become too attenuated to be applicable and allocating costs for harm becomes more complex. Absent a new approach to law and policy, it is unlikely that current liability rules will be sufficient to satisfy the expectations of the judiciary and the public as the underlying technologies develop. Considering these challenges, it is likely that law and policy directed towards AI will require society to accept solutions that may support conflicting values but are beneficial to humanity overall.

I. UNDERSTANDING AI, MACHINE LEARNING, AND THE ROLE OF ALGORITHMS

A. AI, Machine Learning, and Algorithms: A Technical Foundation

In recent years, AI has been thrust to the vanguard of technical development as nation states, private industries, and researchers seek to understand and exploit its potential.[3] Despite its prominence in the global technological realm, there is no universally accepted definition for AI. In a broad sense, AI constitutes a computerized system that can rationally solve complex problems or act appropriately to achieve an objective.[4] Some experts narrow the scope of AI based on taxonomies that reflect the function, capabilities, or problem space of the system.[5] For example, venture capitalist Frank Chen categorizes the problem space of AI into five general groups: logical reasoning, knowledge representation, planning and navigation, natural language processing, and perception.[6] The difficulty in defining what actually constitutes AI stems from the expansive nature of the problems and solutions sought to be conquered through AI, and the underlying performance of algorithms that fuel AI development. Because the problems and

---

[3] Louis Columbus, *McKinsey's State of Machine Learning and AI, 2017*, FORBES (July 9, 2017), https://www.forbes.com/sites/louiscolumbus/2017/07/09/mckinseys-state-of-machine-learning-and-ai-2017/#220930e075b6. ("Tech giants including Baidu and Google spent between $20B to $30B on AI in 2016, with 90% of this spent on R&D and deployment, and 10% on AI acquisitions.")

[4] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 6 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf. ("Others define AI as a system capable of rationally solving complex problems or taking appropriate actions to achieve its goal in whatever real-world circumstances it encounters.")

[5] *Id.* (citing Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (3rd ed. Pearson, 2009)).

[6] *Id.* at 7 (citing Frank Chen, *AI, Deep Learning, and Machine Learning: A Primer* (June 10, 2016), http://a16z.com/2016/06/10/ai-deep-learning-machines).

solutions to be evaluated by AI flow naturally between routine data processing by algorithmic systems and AI machine learning that requires intelligent computer programs, it is common for a problem to be viewed as requiring AI to be solved, but consisting of routine data processing once answered.[7] While the definition of AI may be fluid and inexact, at its core is the pursuit of AI applications that can systemically produce intelligent behavior.[8]

Within the realm of AI and at the center of this technological transformation is machine learning that uses a statistical approach to apply algorithms and learn from data. While AI traditionally involves programming rules and criteria to reach a decision, machine learning uses algorithms to statistically evaluate large amounts of data to repeatedly refine its decision-making processes and outcomes. Machine learning is a subset of AI that involves methods to allow computers to think by "creating mathematical algorithms based on accumulated data."[9] Within machine learning is the concept of deep learning — using neural network models to mimic human thinking. In a sense, deep learning fuels machine learning to create artificial intelligence. Nidhi Chappel, Director of Machine Learning at Intel, described this as "machines learning on their own without explicit programming."[10] She compared this process to how a child observes the world and learns societal norms without being explicitly told the rules.[11]

The inherent ambiguity and breadth of AI in an increasingly blended world of technological advancements and big data has pushed universal definitions of these concepts out of reach, at least for the time being. While precise definitions of these terms are beyond the scope of this paper, it is important to understand that substantively defining AI and its subsets is one of the major hurdles to achieving a practical regulatory context for AI. By defining the technology, the law will be better suited to articulate the problems posed by AI. This is, however, not surprising, as law almost always remains unsettled without concrete examples and practical applications. In this context, the

---

[7] *Id.* at 7 ("In some cases, opinion may shift, meaning that a problem is considered as requiring AI before it has been solved, but once a solution is well known it is considered routing data processing.").

[8] *Id.* ("Although the boundaries of AI can be uncertain and have tended to shift over time, what is important is that a core objective of AI research and applications over the years has been to automate or replicate intelligent behavior.").

[9] Deb Miller Landau, *Artificial Intelligence and Machine Learning: How Computers Learn*, IQ BY INTEL (August 17, 2016), https://iq.intel.com/artificial-intelligence-and-machine-learning/.

[10] *Id.* (quoting Nidhi Chappell).

[11] *Id.*

enthusiasm for investing in AI is great and certain tangible applications have already become pervasive in our daily lives. However, machine learning algorithms — which constitute the core of AI as a vehicle for societal change — remain in infancy, as society has only begun to recognize the potential of AI to transform industries. It is likely that reaching usable definitions of AI and its subcomponents will take time as the challenges to be addressed and the fruits of the technology are realized. In the interim, it is the obligation of technology experts, legal experts, and regulating entities to push through these obfuscated hurdles in pursuit of regulatory definitions that are necessary to shape the legal, social, and political aspects of AI.[12]

B. Promises and Perils: The Next Step in Human Evolution?

Machines that think like humans — a concept that stands as a technological gateway to what could be the next step in human evolution. From autonomous driving vehicles to improving genomic sequencing, AI has just begun to demonstrate its utility in tackling some of the greatest challenges faced by society. On a more intimate level, applications like speech recognition AI platforms are becoming increasingly pervasive in our day-to-day lives. These commercial-oriented forms that have specific applications, known as *Narrow AI*, are becoming an extension of how humans interact with technology to accomplish defined tasks.[13] In a broader sense, *General AI* — a notional future AI system that has the intelligent behavior to process cognitive tasks — is believed to be the future of intelligent systems.[14] While *Narrow AI* allows technology to perform specific tasks beyond what a human can do, *General AI* has the potential to surpass human performance in almost every cognitive task.[15] The capability of AI

---

[12] Defining AI and its subsets is challenging given the technological uncertainties, social considerations, and geopolitical factors that are inherent to a technology with the potential to change how the world operates. In this context, it is likely that defining these terms will take time and will likely require different definitions depending on the context in which the technology is being used.

[13] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 6 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf ("Remarkable progress has been made on what is known as *Narrow AI*, which addresses specific application areas such as playing strategic games, language translation, self-driving vehicles, and image recognition.").

[14] *Id.* ("*General AI* (sometimes called Artificial General Intelligence or AGI) refers to a notional future AI system that exhibits apparently intelligent behavior at least as advanced as a person across the full range of cognitive tasks.")

[15] Max Tegmark, *Benefits & Risks of Artificial Intelligence*, FUTURE OF LIFE INSTITUTE, https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/ ("While narrow AI may outperform humans at whatever its specific task is, like playing chess or solving equations,

systems to evaluate large amounts of complex, and sometimes unrelated, data has the potential to solve some of the world's most enduring problems, leading to enormous social and economic benefits.[16] According to the Center for Data Innovation, AI is already having a positive social and economic impact.[17] The report issued by the Center for Data Innovation highlighted 70 real-world examples of the social and economic benefits in 14 different areas. The Center's director believes that this list has only scratched the surface of AI's social and economic value.

Despite its promises, there are, like most technological innovations, foreseeable and speculative perils. In the short-term, AI has the potential to disrupt job markets across the world. From factory workers to lawyers, AI could replace the need for humans to perform specific tasks. However, the social impact of AI is a matter of debate. While AI may replace humans in specific jobs, many experts believe that it has the potential to create more jobs as technology spurs new industries.[18] From a long-term perspective, AI has sparked concern that it could be a destructive force by either being programmed to do something devastating or to do something beneficial, but with destructive force.[19] The discussion of the practical and theoretical benefits of AI is one that will, and should, continue during the course of the technology's development. However, from a legal context, there are narrower perils that could have a significant impact on society.

AGI would outperform humans at nearly every cognitive task.").

[16] Francesca Rossi, European Parliament Legal Affairs Briefing, *Artificial Intelligence: Potential Benefits and Ethical Considerations* 1, 1 (October 2016), http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/571380/IPOL_BRI%282016%2957 1380_EN.pdf. ("The ability of AI systems to transform vast amounts of complex, ambiguous information into insight has the potential to reveal long-held secrets and help solve some of the world's most enduring problems.").

[17] Daniel Castro and Joshua New, *The Promise of Artificial Intelligence* 1, 2 (October 2016), http://www2.datainnovation.org/2016-promise-of-ai.pdf, (citing Olivia Solon, *Karim the AI Delivers Psychological Support to Syrian Refugees,* THE GUARDIAN (March 22, 2016), https://www.theguardian.com/technol- ogy/2016/mar/22/karim-the-ai-delivers-psychological-support-to-syrian-refugees; Dina Bass, *Microsoft Develops AI to Help Cancer Doctors Find the Right Treatments*, BLOOMBERG, (September 20, 2016), http://www.bloomberg.com/news/articles/2016-09-20/microsoft-develops-ai-to-help-cancer-doctors- nd- the-right-treatments).

[18] *See generally* Douglas Eldridge, *Why the Benefits of Artificial Intelligence Outweigh the Risks*, SMS WIRE (February 22, 2017), https://www.cmswire.com/digital-experience/why-the-benefits-of-artificial-intelligence-outweigh-the-risks/.

[19] Max Tegmark, *Benefits & Risks of Artificial Intelligence*, FUTURE OF LIFE INSTITUTE, https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/ ("when considering how AI might become a risk, experts think two scenarios most likely: 1. The AI is programmed to do something devastating, or 2. the AI is programmed to do something beneficial, but it develops a destructive method for achieving its goal.").

Inherent in AI and machine learning systems is the aspect that the technology's development includes some level of programming and emulation of principles used by human experts. This injects a subjective human element into what is supposed to be an objective and unbiased process that relies on data and mathematical algorithms. But if the data or algorithm is tainted with some level of bias, then how much faith do we have in the AI system? Does an AI system merely create a false expectancy of fairness? No system is perfect, and perhaps these systems, even with a certain level of algorithmic prejudice, are better suited to perform a task in a manner more objective than that of humans. It is likely that the answer to these questions lie with each specific technological use and application. The subjective bias of humans typically cannot be measured and sometimes eludes our own conscious. It is quite conceivable that even with some level of fallacy, AI stands to provide for greater fairness in most decision-making processes. From criminal sentencing to acceptance at an educational institute, AI has the potential to imperfectly perfect human decision-making. But for society to accept the idea that a machine can be fairer than a human requires certain levels of transparency, accountability, and understandability must be accounted for in the technological process that results in a machine that can learn, think, and act in society's best interest.

C. The Regulatory Landscape

1. The United States

In the United States, public and private sectors have placed a heavy focus on the development of AI and machine learning systems. Taking into consideration the short and long-term benefits and concerns of AI, the United States has assessed that the benefits of AI are vast, and the development of AI is critical to the country's economic and social vitality.[20] As such, the United States has devoted significant resources to developing AI and seeks to be at the forefront of AI research.[21] From a regulatory perspective, the United States has assessed that any broad regulation of AI is inappropriate in the current stage of AI and machine learning. As with other technologies that have

---

[20] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 5 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf ("The current and projected benefits of AI technology are large, adding to the Nation's economic vitality and to the productivity and well-being of its people.").

[21] *Id.* ("The United States has been at the forefront of foundational research in AI.").

the potential to significantly impact the global socioeconomic platform, regulation in the early stages of development has the potential to hinder progress in the field. Russian President Vladimir Putin echoed the importance of AI when he stated, "whoever becomes the leader in [AI] will become the ruler of the world."[22] This statement reflects what has been called the new global arms race for superpowers.[23] Thus, it is clear why the United States is not ready to insert a broad regulatory framework into the AI arena. However, this reluctance is not solely based on efforts to win the AI development race, and it does not necessarily mean that the United States is averse to tailored regulation in the future.

The applicability of AI is expansive and already touches on multiple industries that are regulated to protect the public and promote fair competition. The approach of the United States towards regulating AI is based on an informed assessment that accounts for the benefits of the technology, the associated risks to public safety, and potential barriers to innovation.[24] Under this approach, when the risk of an AI-enabled product falls within the realm of an existing regulatory regime, it should be considered whether current regulations already consider the risks in a sufficient manner. If the protections in place are inadequate, existing regulations should be modified and expanded to provide the necessary safeguards while accounting for AI innovation and growth. For example, fully autonomous vehicles are at our doorstep. Rather than seeking to regulate the underlying AI technology, the states and the Federal government are looking to broaden motor vehicle regulations to account for self-driving vehicles. In doing so, the government attempts to minimize barriers to innovation while maintaining its obligation to protect the public.

While the United States government seems to be content to refrain from broadly regulating AI and machine learning for the time being, there is significant debate in the technology industry as to whether this is the right decision. The more alarmist and extreme approach, echoed by technology entrepreneur Elon Musk, urges for government regulation in this field because he believes AI poses an existential

---

[22] *See* Tom Simonite, *For Superpowers, Artificial Intelligences Fuels New Global Arms Race*, WIRED (September 8, 2017), https://www.wired.com/story/for-superpowers-artificial-intelligence-fuels-new-global-arms-race/.

[23] *See generally id.*

[24] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 17 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf ("In doing so, agencies must remain mindful of the fundamental purpose and goals of regulation to safeguard the public good, while creating space for innovation and growth in AI.").

threat to human civilization.[25] However, the United States government believes that even if the more science-fiction threat of AI is true, it does not change how the problem should be addressed today. Essentially, even if AI poses a long-term threat of super-intelligent machines capable of surpassing humanity's control, it does not change how the technology should be pushed forward today. The National Science and Technology Council (NSTC) Committee on Technology has determined that the best way to build the capacity to address these long-term speculations is to tackle the short-term and less extreme security, privacy, and safety risks of AI.[26]

2. The European Union and the Right to an Explanation

Beyond the United States, other democratic nations have taken a more substantive approach to regulating AI. In particular, the European Union passed the General Data Protection Regulation (GDPR) which seeks to, in part, harmonize data privacy laws across Europe. The regulation includes non-discrimination requirements for algorithmic profiling and a right to obtain an explanation of automated decisions that significantly affect users.[27] Bryce Goodman and Seth Flaxman have suggested that this could require "a complete overhaul of standards and widely used algorithmic techniques," and reflects how important human interpretability is to algorithmic design.[28]

While some argue that this regulation is a harmful restriction of AI,[29] Goodman and Flaxman contended that it simply creates an

---

[25] *See generally* Brett Molina, *Musk: Government needs to regulate artificial intelligence* (July 17, 2017), USA TODAY, https://www.usatoday.com/story/tech/talkingtech/2017/07/17/musk-government-needs-regulate-artificial-intelligence/484318001/

[26] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 8 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf ("The NSTC Committee on Technology's assessment is that long-term concerns about super-intelligent General AI should have little impact on current policy. The policies the Federal Government should adopt in the near-to-medium term if these fears are justified are almost exactly the same policies the Federal Government should adopt if they are not justified.").

[27] Regulation (EU) 2016/679 of the European Parliament and Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Recital 71, arts. 13-14, 22, 2016 O.J. L 119/1 [hereinafter GDPR].

[28] Bryce Goodman and Seth Flaxman, *European Union regulations on algorithmic decision-making and a "right to explanation"* 1, 1 (August 31, 2016), https://arxiv.org/pdf/1606.08813.pdf.

[29] *See generally* Nick Wallace, *EU's Right to Explanation: A Harmful Restriction on Artificial Intelligence* (January 25, 2017), http://www.techzone360.com/topics/techzone/articles/2017/01/25/429101-eus-right-

opportunity to ensure transparency and fairness in algorithmic design. The United States has not taken significant regulatory steps towards algorithmic accountability, but it has acknowledged the importance of transparency and fairness in the development of AI.[30] There is a fine line between regulatory actions that will infuse AI and algorithmic machine learning with democratic values of fairness, and those that will create barriers to innovation. At the very least, regulatory efforts to promote transparency and accountability in AI should be explored as the fruits of the technology and the underlying problems posed by innovation continue to develop. The inclusion of these concepts is not only paramount to ensuring justice in a technology that stands to be broadly implemented across industries but may also be critical to social acceptance of AI and algorithmic machine learning applications.

II. NAVIGATING THE FOG OF INNOVATION: THE TECHNOLOGICAL, LEGAL, AND SOCIAL HURDLES TO GOVERNING AI

Beyond the technological hurdles to mastering AI are legal and social challenges. The complexity of these issues is compounded by the fact that AI does not exist in a vacuum. The development and implementation of AI and machine learning is intertwined with data rights, privacy rights, intellectual property protection, ethics, civil rights, due process, geopolitics, and social values, among other considerations. At the core of these issues are the transparency, accountability, and understandability of AI in the judicial system. To ensure fairness and safety, it is expected that AI systems be free of unjustified discrimination and unintended consequences. To promote confidence in these systems, it is necessary that evidentiary mechanisms exist to demonstrate the lack of unintended consequences.[31] Taken a step further, the social acceptance of AI technologies is partially dependent on the existence of legal means to remedy unintended consequences. Part A of this section explores the legal goals and social expectations for a fully integrated AI society, and Part B addresses the problems that stand in the way of reaching these goals and realizing these expectations. Finally, in Part C of this section,

---

explanation-harmful-restriction-artificial-intelligence.htm#

[30] Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence* at 30 (October 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf ("The use of AI to make consequential decisions about people, often replacing decisions made by human actors and institutions, leads to concerns about how to ensure justice, fairness, and accountability.").

[31] See *generally id*. at 30.

the function of judicial decision-making is explored with an emphasis on the significant role that liability litigation is likely to play in shaping the AI landscape.

A. The Discovery of Goals and Expectations: From Legally Simple to Socially Complex

Recent triumphs in AI development have sparked an awakening of the potential utility of AI. Certain AI systems and machine-learned processes have already entered the daily lives of almost all individuals. From the use of AI autopilot systems in commercial flights to the use of AI for facial recognition on social media sites, certain autonomous machine intelligence processes have quietly become a societal norm.[32] While society has yet to experience the scaled infusion of AI into all sectors of our industry, an automated economy fueled by AI may soon be at our doorstep.[33] Across domains, technology continues to provide the seed for imagining new, tangible applications of AI. With each formation of a practical application, society begins to forecast the legal and social considerations of its use and is often faced with contradictory positions. Although these deliberations are typically driven by economics, questions of legal consequences and social expectations underpin any discussion. When considering the scaled application of AI across society, it is necessary to broadly understand the goals of the law and expectations of society.

From a legal perspective, the goals of governing AI are simple — reduce uncertainty and protect consumers without inhibiting innovation. While these concepts are straightforward in idea, they are far from simple in application. From the likelihood of harming the public to providing predictable measures of liability, uncertainty in the law can take many forms. In general, AI governance should be highlighted by some form of validation and verification — making sure an AI system or process does what it is designed to do and providing a mechanism of confirmation to an extent generally accepted by society. However, as discussed later in this paper, while scholars can articulate where the law needs to be in terms of governance, getting there involves complex legal and policy questions. More importantly, there is a cyclical effect in how social expectations drive legal goals and vice-versa. Put simply, social expectations influence how society seeks to

---

[32] *See generally* Gautam Narula, *Everyday Examples of Artificial Intelligence and Machine Learning* (February 14, 2017), https://www.techemergence.com/everyday-examples-of-ai/.

[33] *See generally* Kevin Maney, *How Artificial Intelligence and Robots Will Radically Transform the Economy*, NEWSWEEK (November 30, 2016), http://www.newsweek.com/2016/12/09/robot-economy-artificial-intelligence-jobs-happy-ending-526467.html.

govern AI, but are in-turn influenced by the legal mechanisms in place to account for the unexpected and unjust outcomes that sit on the fringes of technical application.

While policymakers and scholars can speak generally of the broad and lofty legal goals for governing AI, the narrow legal objectives and social expectations are more challenging. Each application of AI does not exist in a vacuum and may be heavily dependent on existential influences — privacy, big data, or security concerns, among others. Society may be willing to accept a specific AI technology, such as self-driving cars, when the benefits to society clearly outweigh the risk of potential harm. However, how does this change when AI is being implemented on a large scale across society? How much deferment to artificial systems is too much, and how does this change as AI becomes scalable on an intimate level? Are we willing to accept artificial intelligence that can drive us to a hospital or even provide a diagnosis, but not a system that decides whether we are eligible for care? In this domain, like many that are at the forefront of innovation, the only certainty is uncertainty. Society may not yet be able to define social expectations for AI. The only guarantee is that large-scale disruption is likely on the horizon, and policymakers will have to consider the legal and social consequences of this transformation.

The social and economic benefits of AI and machine learning processes only go as far as society perceives that the technology is fulfilling its purpose without unjustified bias, prejudice, or harm. As such, the success of this technology is premised on the idea that an AI system must perform its function in a just and fair manner with a minimal and publicly acceptable deviation from absolute expectations. Given the prevalence that algorithms and AI are expected to have in our daily lives, there are certain democratic principles of due process that should be incorporated into algorithmic development to account for a remedy in the instance of harm by an AI system — transparency, accountability, and understandability.[34] However, achieving sufficient levels of justice and fairness in due process is subjective and dependent on the area in which AI is being applied. Society should not expect the same levels of justice and fairness in all domains. Assuring transparency, accountability, and understandability is not an absolute when it comes to AI, it is a spectrum.

How this spectrum is defined moves beyond a legal issue and requires an understanding of social sensibilities as they relate to AI. Despite efforts to perfect an algorithm, the possibility of unintended

---

[34] *See generally* Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 1 (2014).

consequences exist. When AI is used to decide in place of human judgment or to supplement a human decision-making process that has significant consequences, what levels of fairness and accountability are expected to ensure justice? Generally, the level of expected due process is directly proportional to the contiguity of an algorithm to a more fundamental right or critical decision. Put another way, as an algorithm becomes more significant to a decision that by nature requires a higher degree of legal protection, there is a greater need for algorithmic due process. While this may be simple in theory, the complexity of this paradigm makes AI governance an elusive task as the very nature of trying to define the expectations, consequences, and remedies for AI-based harms can easily create multiple and conflicting legal and social positions. Moving forward, the ability of society and the law to conceive mechanisms of accountability for these conflicting values to coexist may define the pace of AI's immersion into humanity.

B. Legal and Social Hurdles to Governing AI Systems

Rapid innovation across various technological domains has compounded the problem of technology outpacing the law. As technology builds on technology, governing mechanisms that are still trying to understand and account for previous innovations struggle to incorporate additional regulatory, ethical, and privacy considerations brought about by "the next big thing." AI and machine learning are no exception. From the Internet to cybersecurity to big data to the Internet of Things (IoT), and now AI, intrinsic to each realm are unresolved considerations surrounding data rights, privacy rights, intellectual property, ethics, due process, social values, and geopolitical concerns. Trying to govern a single technological domain has proven to be difficult enough, yet alone blended worlds of overlapping technologies that complicate regulatory structures at all levels. From local to international governing bodies, the law is lost in the fog of innovation. Meanwhile, societal acceptance and social mores struggle to harmonize the complexity of these technologies that alter the human experience.

The underlying problem of governing AI and related technologies is highlighted by the legal and social trade-offs that have so far eluded regulators. For example, restrictions on big data and privacy can significantly hinder AI systems that rely on available data to perfect their machine-learned processes. Whether it is the "going dark"[35]

[35] *See generally*, James B. Comey, Director, Federal Bureau of Investigation, *Going Dark: Are Technology, Privacy, and Public Safety on a Collision Course?* (October 16, 2014), remarks delivered at the Brookings Institution, *available at* https://www.fbi.gov/news/speeches/going-dark-are-technology-privacy-and-public-safety-on-a-collision-course.

problem or controversy surrounding section 702 of the Foreign Intelligence Surveillance Act (FISA),[36] society struggles to strike a balance between privacy and national security.[37] Moreover, opposing sides tend to take absolute positions that obscure the issue at hand and exacerbate the problem of finding new legal approaches to emerging technologies.[38] For AI, it often seems that the public puts a premium on privacy, while still expecting the benefits of AI systems where the commercial utility is derived from surpluses of consumer data.[39]

Governments at all levels, from local legislators to international regulators, are experiencing the challenge of broadly governing AI technologies.[40] Legislative bodies continue to wrestle with regulating intricate new systems that are rapidly changing and being scaled across industries and society. Uncertainty as to how AI systems will interact with the other complex systems and the economic impact this may have on the development of AI and machine learning systems complicates the question of how to regulate AI. Some existing domains, such as the automobile industry, have more easily embraced some forms of AI regulation.[41] This is perhaps because of the impending transformation of the automobile industry that can no longer be ignored.[42] It is also likely in part because a federal governing body, the National Highway Traffic Safety Administration (NHTSA), already exists to harmonize new technologies with existing laws.

In other areas, however, there is greater concern as to whether existence of any regulation of AI systems would stifle innovation. For example, it is possible that the accuracy and performance of AI systems

---

[36] *See generally* Sneha Indrajit et al., *FISA's Section 702 & the Privacy Conundrum: Surveillance in the U.S. and Globally* (October 25, 2017), https://jsis.washington.edu/news/controversy-comparisons-data-collection-fisas-section-702/.

[37] *See generally* April F. Doss, *Why Changes in Data Science Are Driving a Need For Quantum Law and Policy, and How We Get There*, 14 ABA SCITECH LAWYER 38, 40 (Fall 2017).

[38] *See id.* (discussing how black-and-white views related to national security and privacy obscure the complexity of issues faced by society).

[39] *See id.* at 41 (discussing consumer demands for privacy).

[40] *See generally* Bianca Datta, *Can Government Keep Up with Artificial Intelligence?* (August 10, 2017), *http://www.pbs.org/wgbh/nova/next/tech/ai-government-policy/.*

[41] Gabrielle Orum Hernández, *Interstate Regulatory Alignment: Keys to The Self-Driving Car?* (November 2, 2017), https://www.law.com/corpcounsel/sites/legaltechnews/2017/11/02/interstate-regulatory-alignment-keys-to-the-self-driving-car/ ("Despite most regulatory action at the state level, 62 percent of those polled sought nationally consistent rules from the U.S. Department of Transportation around tech-enabled vehicles.").

[42] Aarian Marshall, *Congress Finally Gets Serious About Regulating Self-Driving Cars*, WIRED (July 19, 2017), https://www.wired.com/story/congress-autonomous-self-driving-car-regulations/ (discussing how emerging technologies in self-driving vehicles is causing parties ranging from tech companies to government watchdogs to agree that the time has come for Congress to regulate autonomous vehicles).

would be jeopardized if engineers were required to prematurely or unnecessarily incorporate functional mechanisms during the developmental process to ensure principles of algorithmic due process — transparency, accountability, and understandability.[43] In contrast, as more trust is placed in AI systems, the lack of these fundamental principles or poor regulatory mechanisms threatens to undermine any benefits offered by AI and machine learning technologies.[44] Any regulatory miscalculation runs the potential of stifling the transformation of industries that are increasingly betting the future on AI technologies.[45] It is not to say that some form of regulation should not be the ultimate goal or that it is even unlikely. Under the Obama administration, the United States began the conversation about the future of AI regulation.[46] However, there is no indication that the United States intends to significantly and broadly regulate AI systems anytime soon.

An alternative to direct regulation of AI systems is the creation of a centralized agency or commission for AI technologies. This would, at the least, allow expertise to be vested to agencies or commissions that can act more quickly than Congress and keep pace with emerging technologies. Any such body would likely be responsible for identifying principles to govern the development and application of AI, as well as enforcing any promulgated standards. In the United Kingdom, calls for an AI commission have been greater than in the United States.[47] However, domestically, the application of AI has yet to warrant the creation of a government entity focused on AI.[48] Rather,

---

[43] Finale Doshi-Velez et al., *Accountability of AU Under the Law: The Role of Explanation* (November 3, 2017), https://arxiv.org/pdf/1711.01134.pdf ("there exist concerns that the engineering challenges surrounding explanation from AI systems would stifle innovation; that explanations might force trade secrets to be revealed; and that explanation would come at the price of system accuracy or other performance objective.").

[44] *Id.*

[45] Manuel Goncalves, *U.S. Auto CEOS More Bullish on AI, Emerging Technology Investments Compared to Global Counterparts: KPMG Survey* (July 27, 2017), https://home.kpmg.com/us/en/home/media/press-releases/2017/07/us-auto-ceos-more-bullish-on-ai-emerging-technology-investments-compared-to-global-counterparts-kpmg-survey.html, (quoting Gary Silberg, Automotive Sector Leader at KPMG LLP, "Our study found higher levels of optimism and confidence by U.S. auto CEOs moving forward in their willingness to invest in emerging technology and innovation compared to their global counterparts.").

[46] *See generally* Ajay Agrawal et al., *The Obama Administration's Roadmap for AI* Policy (December 21, 2016), https://hbr.org/2016/12/the-obama-administrations-roadmap-for-ai-policy.

[47] *See generally* Joel Muckett, *CBI calls for government AI commission* (October 20, 2017), http://economia.icaew.com/en/news/october-2017/cbi-calls-for-ai-commission.

[48] Although the United States has not created an agency or commission devoted to AI, it has created the National Science and Technology Council Subcommittee on Machine Learning and Artificial Intelligence (MLAT). The charter is available at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/ai_charter_-

independent federal agencies have begun to assume the challenge of regulating AI in their respective domains. For example, the United States Food and Drug Administration (FDA) has assembled a team to oversee the use of AI in diagnosing medical conditions.[49] Similarly, the NHTSA has published voluntary guidelines for driverless cars.[50] The lack of federal legislation in the United States and the use of voluntary guidelines by respective agencies may certainly be a sign that the government is relying on private industry and market demands for safe products to drive the development of AI systems.

Even if Congress sought to govern AI through legislation, it is questionable as to whether this is even a task suited for the legislative branch. The tedious and reactionary nature of the legislative process is not matched to keep pace with emerging technologies. Some scholars, such as April F. Doss, argue that our current governing system is insufficient to address the complexity of emerging technologies, and that it is paramount for law and policy mechanisms to evolve to keep pace.[51] While the field is certainly ripe for change, a viable regulatory solution for AI systems and their overlapping technologies continues to elude policymakers and technology experts. In contrast to broad regulation, the more likely approach has been to pursue guidelines and standards that are more flexible and favorable to business investments in AI.

Initiatives in the private industry have emerged, such as those by IBM[52] and nonprofits like AI Global and the Future of Life Institute.[53] Similarly, organizations like the Institute of Electrical and Electronics Engineers have begun to pursue standards in AI and robotics.[54] These standards are not limited to the technical exploration of AI, but include the ethical and moral concerns of AI developers and users. As AI progresses to the point where regulatory standards can be developed, it

---

_signed_final.pdf.

[49] *See generally* Jeremy Hsu, *FDA Assembles Team to Oversee AI Revolution in Health* (May 29, 2017), https://spectrum.ieee.org/the-human-os/biomedical/devices/fda-assembles-team-to-oversee-ai-revolution-in-health.

[50] *See generally* Colin Dwyer, *Department of Transportation Rolls Out New Guidelines for Self-Driving Cars*, NPR (September 12, 2017), https://www.npr.org/sections/thetwo-way/2017/09/12/550533833/department-of-transportation-rolls-out-new-guidelines-for-self-driving-cars.

[51] *See generally* April F. Doss, *Why Changes in Data Science Are Driving a Need For Quantum Law and Policy, and How We Get There*, 14 ABA SCITECH LAWYER 38 (Fall 2017).

[51] *Id.*

[52] Kay Firth-Butterfield, *Artificial Intelligence and the Law*, 14 ABA SCITECH LAWYER 28, 30 (Fall 2017) (citing *Transparency and Trust in the Cognitive Era,* IBM THINK Blog (Jan 17, 2017).

[53] *See id.* at 30.

[54] *Id.*

is likely that these standards will be derived from best practices promulgated in a manner similar to the National Institute of Science and Technology's (NIST) cybersecurity guidelines.[55] In the absence of any regulation of AI on the horizon, set standards would likely have a significant impact on the AI landscape in the United States. Through standardized approaches to developing and testing AI systems, the United States is moving to an industry and liability approach to governing AI that has the potential to vest the judicial system with significant influence in shaping the future of AI.

C. Judicial Decision-Making as a Driving Force in Governance

AI and the technological domains it touches have proven to be problematic for direct regulation. From data privacy to cybersecurity, regulation has been a fragmented process that eludes traditional vehicles of law and policy. Absent unequivocal regulation, transforming industries seeking to capitalize on the benefits of AI have begun to look towards an industry-driven approach where guidelines and best practices provide some level of security to direct investment. Arguably, this approach will provide greater adaptability and flexibility than direct legislation in keeping pace with emerging and rapidly developing technologies. It is not to say that this approach is ideal. The business interests of private industry may not be aligned with those of society as a whole and may even contradict public safety concerns. As the sentiments of Ms. Doss were previously conveyed, increasingly complex issues in technology, privacy, and security foster a need for a new approach in law and policy to counter the uncertainty of emerging technologies.[56] However, until a new way of thinking takes hold, standards, guidelines, and best practices are likely to rule the AI world, and judicial decision-making will follow right behind.

Assigning and quantifying liability for actions taken or influenced by AI systems will play a significant role in the maturation of the AI landscape. With the development of standards and best practices in AI development, testing, and application, AI will fall into the purview of the court system through tort liability. While the intricacies of liability in a world of AI are uncertain and stand in the early stages of exploration, as discussed in Section III, standardized approaches to AI and machine learning systems will require the courts to weigh in. On a

---

[55] *Id.* (suggesting that regulatory standards can be built on existing ones, such as the U.S. National Institute of Standards and Technology (NIST) standards for cryptography).

[56] April F. Doss, *Why Changes in Data Science Are Driving a Need For Quantum Law and Policy, and How We Get There*, 14 ABA SCITECH LAWYER 38, 42 (Fall, 2017).

case-by-case basis, the judicial system will slowly serve as an indirect form of regulation through the development of legal standards, precedence, and subsequent deterrence. However, there are inherent fallacies that cast doubt on the effectiveness of using the judicial system to indirectly regulate AI.

A serious weakness within the judicial system is the lack of technological expertise. Few judges and lawyers have sufficient technological training to fully understand and judicially account for complex technologies that have already been brought before the courts. Several legal scholars have proffered that it is incumbent on the legal community to teach certain aspects in law school to help prepare legal enterprises to tackle the issues that lie ahead.[57] However, this long-term solution does not provide immediate relief for the technological storm at the shores of the judiciary system. As Chief Justice Roberts suggested, "[judges] haven't yet really absorbed how [AI] is going to change the way we do business."[58] Another significant limit of the judiciary is the remedial nature of courts. Courts are designed to correct or compensate for harm that has already occurred. Although legal precedence will certainly have a deterring effect, the process of establishing legal norms may have difficulty keeping up with the accelerated pace of AI development given the lengthy litigation process that is common in legal cases surrounding complex technologies.[59] A final drawback of the judiciary system to indirectly govern AI is the court's narrow focus.[60] Judicial exploration is typically limited to specific facts relevant to the harm and risk involved and does not typically account for broader considerations surrounding the social and economic benefits of emerging technologies.[61]

Despite the limitations of judicial decision-making to influence the governance of AI, there are strengths in the judiciary that make it suited to be a driving force in governing emerging technologies. Although the judicial system may lack technical knowledge, courts have significant expertise in allocating responsibility. In the world of

---

[57] For example, *see* April F. Doss, *Why Changes in Data Science Are Driving a Need For Quantum Law and Policy, and How We Get There*, 14 ABA SCITECH LAWYER 38, 42 (Fall 2017).
[58] *See* Adam Liptak, *Sent to Prison by a Software Program's Secret Algorithms*, THE NEW YORK TIMES (May 1, 2017), https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html.
[59] Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J. L. & TECH. 353, 388 (2016) ("Tort law influences future behavior primarily through the deterrent effect of liability.").
[60] *Id*. ("Once a suit is brought, procedural and evidentiary rules act to focus attention on the specific facts that led to harm in that case; the ability to introduce information regarding broader social and economic considerations is limited.").
[61] *Id*.

AI and autonomy, there can be numerous entities involved in the development and implementation of an AI product that results in harm, and the courts are perhaps best equipped to allocate responsibility across an enterprise of defendants.[62] Further, when uncertainty is high, the legal fact-finding processes of the tort system can be a dominant force in developing specific and relevant information when causation is in question.[63] While at times the litigation process may be tedious and adversarial to the detriment of the public as a whole, it does provide for a natural, albeit slow, development of the law where workable solutions are promulgated across jurisdictions and harmful laws are rejected or modified.[64]

These benefits demonstrate that deferring to judicial decision-making to indirectly govern AI systems and emerging technologies is not completely the result of a lack of a better alternative. Rather, it is a reflection of a fragmented legal and policy process that struggles to absorb complex and esoteric technologies scaled across industries and society. AI may be at the world's doorstep, but full autonomy in these technologies has not yet warranted rash action to curb systems that may prove to have significant benefits to society. While the capacity of AI as a tool for humanity seems to become more transparent day by day, legal and social regimes can only speculate as to the latent benefits to society that fully autonomous systems may hold. It is, therefore, quite possible that the ability of current liability regimes to address the questions posed by current AI technologies -- those that are more useful than revolutionary -- may pave the way for how society manages fully autonomous and truly intelligent AI that may be yet to come.

III. TRADITIONAL LIABILITY REGIMES AND AI SYSTEMS

As the spectrum of intelligent machines, from *General AI* to fully autonomous and intelligence systems capable of independent learning, continues to evolve, the judiciary will be required to confront increasingly complex issues of liability. Litigation surrounding harms caused by automated machines and AI systems has already entered the

---

[62] *Id*. at 389 ("Because courts have more experience than the other institutions in allocating responsibility in such situations, they remain best equipped to make such determinations of responsibility when harm occurs.").

[63] *Id*. ("The intensive discovery and fact-finding processes of civil litigation provide powerful tools for unearthing relevant information regarding the design and safety features of a harm causing product, and gaining such specific and detailed information is particularly important when uncertainty regarding causal factors is high.").

[64] *Id*. at 391 ("the incremental nature of the common law provides a mechanism that allows legal rules to develop organically;").

courts purview.[65] However, as the underlying technologies push towards greater autonomy, the limits of common law tort regimes will be tested as the application of new technologies to traditional theories of liability becomes increasingly convoluted. How well the judiciary responds and applies current legal frameworks to effectively remedy harms is likely to play a significant role in the future of AI governance as the technology pushes towards greater autonomy. This section explores the application of liability claims involving AI systems to common law tort regimes.

Part A introduces the relevance of legal identity as it relates to AI technologies by examining how the legal identity of an AI system plays a significant role in the application of common law tort claims. Part A also explores the idea that as AI systems assume more autonomy, traditional liability claims will be more strained in addressing respective harms. In Part B, current legal frameworks of liability are assessed as applied to harm caused by AI systems. Finally, Part C suggests that there is no common solution to the liability problem of AI systems, and that the law will likely require a blended approach to address the legal challenges on the horizon. The ability of current law and policy regimes will not be able to solve, but only manage, the novel legal issues that arise as technology pushes the envelope of autonomy and machine intelligence. Eventually, the scientific advancements that seek to enhance the human experience will likely require an evolution of law and policy.

A. Identifying the Defendant: Principal-Agent vs. Alternate Legal Identities

How an AI system is perceived and the legal status it is afforded is fundamental to resolving issues of liability for harms caused by autonomous systems. When the mark of human contribution is apparent in the decision of an autonomous system, common law principles of liability may be sufficient to find fault.[66] In instances of an identifiable defect in an AI system, product or manufacturer liability may vest responsibility for the harm to the developers or those involved in the production chain of an AI system. Furthermore, under the principal-agent concept, an AI system could be considered an agent of

---

[65] *See generally Artificial Intelligence Litigation: Can the Law Keep Pace with the Rise of the Machines?* (December 2016), https://www.quinnemanuel.com/the-firm/publications/article-december-2016-artificial-intelligence-litigation-can-the-law-keep-pace-with-the-rise-of-the-machines/.

[66] David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence* 89 WASH. L. REV. 117, 120 (2014) ("Where the hand of human involvement in machine decision-making is so evident, there is no need to reexamine liability rules.").

a manufacturer or other entity that, in some form or another, directs or is responsible for the harm caused. In this instance, when the law can point to some discernible level of human involvement, common law tort regimes may be suited to remedy a situation. However, as discussed in Part B of this section, the applicability of traditional liability regimes to AI is not clear and may likely be deficient to adequately protect the public when complex technologies are involved.

As AI systems move towards greater autonomy, existing liability rules are likely to become insufficient for assigning fault for harms caused by AI technologies. A fully independent autonomous machine — one that is so far removed from human control that it acts based on its own analysis and without direct human input — certainly brings into question the principal-agent concept. In this instance, does such human deferment to a machine break any connection to the manufacturer or creators of the underlying algorithms to the extent that a machine could no longer be viewed as an agent of a principal? If so, to whom, or what, should fault be accorded? Perhaps a failure to introduce appropriate data in the machine learning process or failing to reasonably forecast a potential outcome could pivot the spotlight of responsibility to an entity in the production chain. But what about when fault is unclear, cannot be determined, or is the result of an unforeseen harm? Does the concept of *res ipsa loquitur* apply? These questions reflect the anticipated strain that AI is expected to place on common law principles of liability.

Alternatives to making liability determinations based on the agency of an AI system have been offered and revolve around different concepts of legal identity. Over twenty-five years ago, Lawrence Solum explored the idea of whether AI could become a legal person.[67] At the time, Solum believed this inquiry was only theoretical because the technology to justify judicial review of the matter was lacking.[68] However, the recent emergence of more autonomous AI systems has shown that the time for this judicial inquiry may soon be here. While the notion of treating AI as a person may seem unnatural, the concept of juridical persons that are recognized in law as having duties and rights of those of natural persons shows that the idea of endowing machines with the legal status of a person is not so far-fetched.[69] After

---

[67] *See generally* Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV 1231 (April 1992).

[68] Cleve R. Wootson Jr., *Saudi Arabia, which denies equal rights, makes a robot a citizen*, THE WASHINGTON POST (October 29, 2017), https://www.washingtonpost.com/news/innovations/wp/2017/10/29/saudi-arabia-which-denies-women-equal-rights-makes-a-robot-a-citizen/?utm_term=.f491ddf747c0.

[69] *Juridical Person*, BLACK'S LAW DICTIONARY (10th ed. 2014).

all, this bestowment has already occurred in numerous instances for corporations.[70]

The idea of some form of corporate personhood for individual AI systems could make these technologies civilly or criminally liable for their actions. However, under current thinking, even if such a status was granted in the eyes of the law, "robots cannot be sued."[71] Furthermore, in the instance of a liable robot defendant, compensation would still be required at a corporate level. While this would likely occur in some form of insurance either individually or as a pool, obligations for compensation could be passed to the owner of the system where the owner is strictly liable for the harm of the AI system.[72] Similarly, an AI system could be treated as an employee and the owner as an employer. In this instance, the owner may be required to ensure some adequate level of supervision.[73]

The question of legal identity is not straightforward and will likely be complicated by the nuances of specific technologies. The only certainty in this realm is that shifts to greater autonomy in AI will continue to drive a need to, at the very least, reassess the capacity of traditional liability regimes to provide just compensation for those harmed. Until technology forces the judiciary to adopt new frameworks of legal identity for AI systems, it is likely that litigation will be driven by the view that principals, such as developers, manufacturers, and owners, are directly responsible for the AI they choose to implement. As traditional forms of liability are increasingly tested, new legal questions will be raised that will result in justified uncertainty as to how adept the judiciary system is to meet the liability challenges brought forth by AI.

B. Applying Traditional Liability Regimes

With the rise of autonomy in AI systems, the legal community and society will be forced to consider novel questions of liability surrounding wrongful acts made or influenced by AI. Where a sufficient connection is apparent between harm caused by an AI system

---

[70] *See generally* Ciara Torres-Spelliscy, *The History of Corporate Personhood* (April 7, 2014), https://www.brennancenter.org/blog/hobby-lobby-argument.

[71] David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence* 89 WASH. L. REV. 117, 124 (2014) (citing United States v. Athlone Indus., Inc., 746 F.2d 977, 979 (3d Cir. 1984)).

[72] Matt Scherer, *Digital Analogues (Intro): Artificial Intelligence Systems Should Be Treated Like…*, FUTURE OF LIFE INSTITUTE (June 9, 2016), https://futureoflife.org/2016/06/09/digital-analogues-intro-artificial-intelligence-systems-treated-like/ (discussing corporate personhood for AI systems).

[73] *Id.*

and the underlying human contribution, legal questions of responsibility are more evident. In instances of a programming error or identifiable manufacturing defect, fault may be clear. However, as technology becomes more complex and convoluted, whether it be more lines of code or more entities involved in the production process, traditional avenues of assigning responsibility will become less apparent. This sub-section explores traditional theories of liability for determining fault as applied to the farther end of the AI spectrum -- when autonomy is significant, and culpability cannot be easily discerned.

Product liability as a legal framework is perhaps the most acceptable form of assigning responsibility for harm arising from AI technologies. It typically involves claims surrounding a manufacturing defect, design defect, information defect, or a failure to warn. From flawed programming to using low quality data during the machine learning process to failing to fully warn a consumer of dangerous consequences, AI can certainly fall within the purview of product liability. However, product liability is premised on the idea that fault is discernible. For example, if a plane crashes because of an error in the autopilot system, the law will likely hold the developer of the system liable either directly or through indemnification. In this instance, the legal analysis falls squarely within the realm of product liability in some form or another.[74] Conversely, it is not to say that this analysis would be straightforward. AI technologies are complex, so it is likely that litigation may be muddled by a fact-finding process laced with finger pointing and compounded questions of blame.

The courts were recently confronted with an issue that will likely challenge AI claims in a similar fashion — whether an auto manufacturer can be held liable for a defect when the presence of a defect can be inferred, but not identified. In a case involving vehicles manufactured by Toyota, a suspected software defect caused a sudden acceleration that could not be stopped by the driver.[75] In these cases, engineers were unable to identify a design or manufacturing flaw that caused the sudden acceleration.[76] Although the plaintiffs could not isolate a specific defect, the court found that there was sufficient evidence to which a reasonable jury could infer that the accident was

---

[74] For a more in-depth analysis regarding the application of types of product liability for autonomous vehicles, *see* David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence* 89 WASH. L. REV. 117, 129 (2014).

[75] *See* In re: Gen. Motors LLC Ignition Switch Litig., No. 14-MD-2543 (JMF), 2015 WL 9480448 (S.D.N.Y. Dec. 29, 2015); *see also* In re Toyota Motor Corp. Unintended Acceleration Mktg., Sales Practices, & Prod. Liab. Litig., 978 F. Supp. 2d 1053 (C.D. Cal. 2013).

[76] *Id.*

more than likely caused by the car and not the driver.[77] Given the substantial number of complaints and commonalities among them, Toyota has sought to settle cases for economic and personal injury claims.[78]

The more difficult consideration is whether product liability is sufficient to correct harm when it cannot be reasonably inferred that a defect contributed to the injury. When an AI machine begins to deviate from its programmed priorities, a sign of true autonomy, agency principles become less relevant.[79] This bares the questions, what is it about autonomy that can render it defective within the meaning of the law and how do we manage these risks? When there is injury without a discernible flaw, defect, or failure arising from human contribution, who should bare the loss? Traditional concepts of product liability will likely fail to provide relief because a manufacturing defect cannot be identified, and the reasonableness of a jury to infer the cause of the harm will become increasingly attenuated.

Absent direct evidence of fault, the law has traditionally looked towards the doctrine of *res ipsa loquitur*. The idea that the very nature of harm infers negligence, however, does not resolve the questions of liability for fully autonomous systems. Under *res ipsa loquitur*, a defendant can negate any inference of the necessary elements of duty of care, breach, and causation by an evidential showing that the defendant's conduct was not negligent.[80] More importantly, this doctrine surrounds the inference that someone was a fault. If the harm in question is unexplainable, untraceable, and rare, then the elements of *res ipsa loquitur* likely cannot be satisfied. This doctrine may, however, survive a determination of fault when an injury is not rare and a commonality exists between potential plaintiffs.[81] In this instance, similar to the Toyota litigation, it may be easier for a jury to infer that something must have gone wrong to cause the injuries in question.

---

[77] *See* In re Toyota Motor Corp. Unintended Acceleration Mktg., Sales Practices, & Prod. Liab. Litig., 978 F. Supp. 2d 1101 (C.D. Cal. 2013).

[78] David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence* 89 WASH. L. REV. 117, 142 (2014) (discussing settlements in the Toyota litigation).

[79] *See id.* at 145 ("A machine that can define its own path, make its own decisions, and set its own priorities may become something other than an agent.").

[80] J. D LEE and BARRY A. LINDAHL, MODERN TORT LAW: LIABILITY AND LITIGATION § 15:25 (2d ed. 2008).

[81] David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence* 89 WASH. L. REV. 117, 143 (2014); Estate of Edward W. Knoster v. Ford Motor Co., 200 F. App'x 106, 114 (3d Cir. 2006) (applying New Jersey law and finding that section three of the *Restatement (Third) of Torts* preserved the res ipsa loquitur inference so that sometimes when a product fails, "'common experience' indicates it would not have done so absent a defect").

Some scholars, such as Professor of Law David C. Vladeck, have proposed that determining fault must be removed from the equation when a failure cannot be reasonably attributed to the product.[82] In terms of self-driving cars, Professor Vladeck contends that a strict liability regime that deviates from standard negligence claims is required because the technologically complex nature of these systems can make it impossible for an injured party to overcome the traditional defenses to negligence.[83] This position is also premised on the idea that the advanced capacity of these systems is so great that they are not expected to fail.[84] Where the benefits to society of an autonomous technology are so abundant and the risks so rare, a strict liability regime may in some cases be sufficient to account for the fringe cases of harm caused by an AI system. At the very least, it would likely reduce uncertainty in support of innovation.

Taken a step further, it is possible that a strict liability system would have a deterring effect that would serve the goals of an AI-integrated society when the risk-benefit dichotomy is not so clear. Developers and manufacturers of autonomous AI systems would be discouraged from taking a product to market that is lacking in effective safeguards. While this may stifle innovation in the sense that it would slow the pace in which a product reaches consumers, companies seeking to reduce uncertainty may be influenced to maximize utility and minimize risk before sale. Should science prove correct, driverless vehicles will be a prime example where the benefits to society so outweigh the risks of rare and unexplainable harms that those in the production process can predictably absorb the costs of harm.[85] Similarly, a strict liability regime may indirectly push developers of other technologies who seek to capitalize on the economic benefits of implementing AI into their industries towards a similar cost-benefit scenario that is mutually beneficial for industry and society.

With strict liability comes the task of appropriating damages, who should bear the costs of harm? The most likely entity to be responsible would be the manufacturer because the manufacturer is the entity that typically controls how, when, and at what cost an AI system enters the

---

[82] David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence* 89 WASH. L. REV. 117, 146 (2014) ("My proposal is to construct a system of strict liability, completely uncoupled from notions of fault for this select group of cases.").

[83] *Id.* at 146. (discussing strong policy reasons in support of a strict liability regime for self-driving vehicles).

[84] *Id.* at 146 ("it is precisely because these machines are so technologically advanced that we expect them not to fail.").

[85] *Id.* at 146 (discussing the suitability of vehicle creators to absorb costs as compared to the injured party).

market. In doing so, the manufacturer can offset liability costs by incorporating some form of an insurance premium into the sale price.[86] However, a narrow view of allocating liability costs to the manufacturer can be flawed. AI technologies are complex and likely to have many "hands in the pot" as a product goes to market. Given the inherent problem of determining fault among the numerous programmers and developers, it could very likely be that the undeterminable fault lies with one or more entities in the production chain. Therefore, it may be more appropriate to spread the costs across the companies involved.[87] More importantly, placing monetary responsibility solely on the manufacturer contrasts with the overall goal of public safety. Insulating those that produce a component of an AI system may encourage product development that sacrifices the public interest for economic benefits beholden only to a single entity.

In the context addressed so far, strict liability has been considered for cases such as self-driving vehicles where the AI technology was directly involved in the harm. But it is also necessary to consider how allocating costs under a strict liability regime fares when the connection between the technology and the victim is separated by some degree. This problem can be envisioned by considering the use of AI in professional judgment. If an autonomous AI system — one that learns, thinks, and acts absent of human control — makes a recommendation, and a medical doctor acts based on this information to the detriment of a patient, should those in the production chain still bare some cost of the injury even if the outcome was highly unanticipated? This scenario creates a litany of issues, such as the duty of care a medical provider owes to his or her patient when AI technologies are involved in treatment.

This issue will become even more complex when externalities, such as insurance coverage, come into play. Can a doctor forego the recommended treatment of an AI system in favor of a care that is more complimentary to insurance coverage? If the doctor chooses to make insurance coverage a priority, are the entities that helped create the technology absolved from paying damages even though the recommendation of the system was unrealistic given the totality of the circumstances. These dilemmas only scratch the surface of legal issues that are apt to arise as AI becomes infused into more and more domains.

---

[86] *Id.* at 148 ("because the manufacturer sets the price for the vehicle, and so the manufacturer can build in an "insurance premium" into the vehicle's sale price to offset expected liability costs.").

[87] *Id.* at 148 (discussing concerns associated with making a manufacturer bare all costs of liability versus a cost-spreading approach that requires suppliers of the final product to absorb some costs of liability).

More importantly, they demonstrate how current liability regimes will be strained to account for the nuances of AI technologies. While a strict liability regime may be a broad solution to the liability of AI systems, the law will likely struggle to apply these concepts to AI across the board.

C. The Long-Term Reality: Narrowing the Gap Between Law and Technology

Traditional mechanisms of determining fault and assigning costs to compensate those harmed by AI are expected to vex current liability regimes as the technology pushes towards fully autonomy. Absent a new approach to law and policy, it is unlikely that current liability rules will be sufficient to satisfy the expectations of the judiciary and the public. With each technology, industry, and application, there will exist nuances that cannot be accounted for under common-law tort regimes. This problem faced by society, a lack of a uniform solution to the liability of AI, will likely require a blended approach that incorporates various elements of tort law. AI does not exist in a vacuum, nor will its solutions. Society will likely be required to accept remedies that may, at times, be laced with conflicting values but are largely beneficial to humanity. Liability is just one consideration in the multifaceted world of AI and governance. One could argue that AI is so complex that specific technologies cannot be effectively governed across the board; rather, society's approach must be geared towards governing specific applications and outcomes. Across the world of innovation, the law struggles to keep pace. How society narrows the gap, if at all, will require legal ingenuity on par with the same technical innovations that spawn a new world where the human experience is augmented by machines.

CONCLUSION

For AI in the law, the challenges are great, the scope is vast, and the implications are significant. In the realm of imagination, AI is a spectrum of utility ranging from mundane tasks to the ability to think, learn, and act in a manner more intelligent than humans. Although AI technology stands in a relative infant state, scientific breakthroughs are rapidly shifting specific applications of AI from a conceptual dream to a tangible reality. With each concrete realization, the debate over the utility of AI is invigorated with questions of law and social policy. As with other emerging technologies, AI will continue to push for new approaches in law and policy as it stands to disrupt normal notions of governance. In the near-term, AI technologies will likely elude

regulatory efforts and instead be subject to case-by-case determinations of liability in civil courts. However, as the technology moves towards greater autonomy, it will become increasingly more difficult to harmonize the nuances of innovation with traditional concepts of law. Moving forward, to capitalize on the benefits offered by AI, or more importantly to avoid its destructive perils, it will be paramount for society and the law to evolve.