Penn State Journal of Law & International Affairs

Volume 9 | Issue 1

February 2021

The Genie Is Out Of The Bottle: What Do We Wish For The Future Of AI?

Kelly Carman

ISSN: 2168-7951

Penn State
Journal of Law & International Affairs

2020 VOLUME 9 NO. 1

THE GENIE IS OUT OF THE BOTTLE: WHAT DO WE WISH FOR THE FUTURE OF AI?

By Kelly Carman[*]

[*] Kelly Carman is an Associate Editor of The Journal of Law and International Affairs and a 2021 Juris Doctor Candidate at The Pennsylvania State University School of Law.

I. INTRODUCTION

A. Artificial Intelligence in Everyday Life

"Alexa, what's the weather?" Chances are, we have all been there. Sure, we could walk outside and feel the weather for ourselves, but the combination of the convenience and the futuristic feeling of receiving an accurate response out of thin air has led to the widespread popularity of the countless Alexa-enabled devices we use today. The origin of this technology is interesting to think about. What ignited the fire behind the creation of the smart assistant we have come to know and love? Turns out, Alexa's beginnings are more humble than you may have imagined. According to David Limp, Amazon's senior vice president of devices and services, Alexa was inspired by science fiction.[1] Limp recounted that Alexa creators "aimed to reinvent the conversational computer onboard *Star Trek's* Starship Enterprise."[2] The Starship Enterprise employs the Library Computer Access and Retrieval System ("LCARS") to operate the computers on each spacecraft and in the facilities.[3] Not unlike Alexa, who both responds to a user's natural voice and can be controlled through the Alexa mobile application, LCARS can be controlled by keypad commands as well as voice commands.[4] A more in depth description of the capabilities of LCARS is as follows:

> LCARS used sophisticated subroutines in order to understand and execute vocal natural language commands. This enabled even complicated tasks to be executed with just a few commands in the case of voice or button presses in the case of keypad commands. LCARS controlled the retrieval and storage of files in the data banks housed within the

---

[1] Julie Bort, *Amazon engineers had one good reason and one geeky reason for choosing the name Alexa*, BUSINESS INSIDER (Jul. 12, 2016, 7:05 PM), https://www.businessinsider.com/why-amazon-called-it-alexa-2016-7.

[2] *Id.*

[3] *Library Computer Access and Retrieval System*, MEMORY ALPHA, https://memory-alpha.fandom.com/wiki/Library_Computer_Access_and_Retrieval_System (last visited Jan. 12, 2020).

[4] *Id.*

> ship's computer cores including logs like personal logs or transporter logs[.] . . . It was used to retrieve files in external databases[,][5] . . . display recently recorded data[,][6] . . . [and display] results of analyses . . . [.][7][8]

Comparing these abilities of *Star Trek*'s LCARS alongside the capabilities of Alexa seems to suggest the developers were successful in their desire to design the innerworkings of Alexa essentially as a working model of LCARS. As for the name itself, Alexa's name is reminiscent of Egypt's Ancient Library of Alexandria.[9] The Library of Alexandria[10] is said to have been envisioned "as a library that would house a copy of every book in the world, an institution to rival those of Athens itself."[11] With this vision in mind, it makes sense that Amazon wished for users to see Alexa the same way: with the ability to harness all of the knowledge in the world, as something that would rival all technology that came before it.

Finally ready for the debut, Amazon launched its Alexa voice service for the first time on the Amazon Echo in November of 2014.[12] Alexa is Amazon's cloud-based voice service that gives users the opportunity to interact with everyday technology.[13] Alexa is a

---

5 *See id.* (*e.g.*, when a ship needed to assess traffic control around bases or investigate inventory databases of surplus depots).

6 *See id.* (*e.g.* medical scans, tactical scans and sensor scans).

7 *See id.* (*e.g.*, linguistic analyses, configuration analyses of system networks, and magnetic flux density analyses).

8 *Id.*

9 Bort, *supra* note 1.

10 *See* Brian Haughton, *What happened to the Great Library at Alexandria?*, ANCIENT HISTORY ENCYCLOPEDIA (Feb. 1, 2011), https://www.ancient.eu/article/207/what-happened-to-the-great-library-at-alexandria/. The Library of Alexandria was once the largest library in the ancient world. While the origins of the Library of Alexandria are not fully clear, it is believed that around 295 BCE, the scholar and orator Demetrius of Phalerum, an exiled governor of Athens, convinced Ptolemy I Soter to establish the library.

11 *Id.*

12 *See* Brandon Vigliarolo, *Amazon Alexa: Cheat sheet*, TECHREPUBLIC (Sept. 27, 2019, 3:27 AM), https://www.cnet.com/news/amazon-sold-tens-of-millions-of-echo-devices-in-2018/.

13 *Id.*

voice-controlled virtual assistant, who performs tasks through the use of "skills".[14] Alexa comes with many skills included, but you can download many additional ones on your own.[15] While most people know what the Alexa voice service is, and possibly even own an Alexa-enabled device, chances are most people do not know what gives Alexa the ability to uncover the answers to its users' questions. How many of you knew that by opening your home to Alexa, you were also welcoming in artificial intelligence ("AI")?

B. How Common is AI?

Many have questioned just how widespread Alexa has become, and at the conclusion of 2018, Amazon finally gave us some answers. After years of dodging the question, Amazon shed light upon how many Alexa-enabled devices had actually been sold since the voice service's launch four years prior. David Limp revealed that more than 100 million Alexa-enabled devices had been sold so far.[16] Limp did not break down specifics on how many of each device had been sold, but he highlighted that tens of millions of Echo devices were sold and the number of people who interact with Alexa daily doubled from the previous year.[17] The year-end statistics also uncovered that, by 2018, Amazon's developer community had built more than 70,000 skills for Alexa, and increased the number of Alexa compatible smart home devices to more than 28,000 devices from more than 4,500 different brands.[18] Further, customers were using

---

[14] *See* Taylor Martin, *The 50 most useful Alexa skills*, COMPUTER NETWORK (Feb. 9, 2019, 4:00 AM), https://www.cnet.com/how-to/amazon-echo-most-useful-alexa-skills/. Automatic can track the location of your car; Harmony can control your entertainment system; Meal Idea will provide you with recipes; Big Sky can give you the weather forecast for a specific address; Short Bedtime Story allows you to custom add your children's names into the bedtime stories it will tell; and Lyft allows you to ask pricing for pricing of specific rides.

[15] Vigliarolo, *supra* note 12.

[16] Toni Reid, *Everything Alexa Learned in 2018*, THE AMAZON BLOG DAY ONE (Dec. 19, 2018), https://blog.aboutamazon.com/devices/everything-alexa-learned-in-2018.

[17] Gordon Gottsegen, *Amazon Sold Tens of Millions of Echo Devices in 2018*, COMPUTER NETWORK NEWS (Dec. 19, 2018, 12:18 PM), https://www.cnet.com/news/amazon-sold-tens-of-millions-of-echo-devices-in-2018/.

[18] Reid, *supra* note 16.

Alexa millions of times each day. Developers made improvements[19] to the Alexa software to make interacting with Alexa more natural.[20] Alexa also became more popular in the world of visual entertainment, by being integrated into Fire TV and Fire tablets.[21] If these advancements and accomplishments were not impressive enough, Alexa also told over 100 million jokes in 2018 alone.

Consumers seem to love Alexa, and there are studies to prove it. In September of 2018, Nielsen Global Connect,[22] powered by The Nielsen Corporation,[23] announced the launch of the Nielsen MediaTech Trender.[24] The MediaTech Trender is a consumer tracking survey focusing on perceptions of emerging technological

---

[19] *See* Day One Staff, *How Alexa keeps getting smarter*, THE AMAZON BLOG DAY ONE (Oct. 10, 2018), https://blog.aboutamazon.com/devices/how-alexa-keeps-getting-smarter. New context carryover technology allows Alexa to track references through multiple rounds of conversation, and natural skill interaction erases the need for users to specifically name each skill they wish to use and instead helps Alexa select the skill that corresponds with the user's general request.

[20] Reid, *supra* note 16. Alexa Skill Blueprints allowed customers to create their own personalized skills for Alexa through templates. Customers are using Skill Blueprints to keep track of household chores, leave personalized instructions for houseguests, and more. Additionally, developers have added Multi-step Requests, Follow-Up Mode, and Context Carryover to Alexa's Skillset. But one of the most interesting features was Whisper mode, which allows Alexa to detect if you are whispering to her and adapt her responses to match.

[21] *Id.* Customers are able to listen to music on Alexa with support for Amazon Music, Apple Music, Tidal, Deezer, Pandora Premium, and more. Additionally, Alexa allows customers to stream or watch live TV through integrations with Fire TV, Xbox, TiVo, Dish, Netflix, and DIRECTV.

[22] *See Nielsen Global Connect*, NIELSEN, https://www.nielsen.com/us/en/solutions/nielsen-global-connect (last visited Jan. 12, 2020). Nielsen Global Connect provides open measurement and analytics platforms for retailers and manufacturers. The company hopes hopes to drive progress in the digital and technological markets by encouraging manufacturers and retailers to make bold decisions and transform their businesses with trusted data, solutions and insights.

[23] *Id.* The Nielsen Corporation is a global provider of market research and analysis of media and viewer interaction.

[24] *Nielsen Launches New MediaTech Trender Survey to Uncover Consumer Sentiment on Emerging Technologies*, NIELSEN (Sept. 27, 2018), https://www.nielsen.com/us/en/press-releases/2018/nielsen-launches-new-mediatech-trender-survey-to-uncover-consumer-sentiment-on-emerging-technology/.

devices.[25] The goal of this survey was to gather consumers' opinions surrounding up and coming technologies, such as smart speakers.[26] The survey was conducted among 2,000 United States consumers, aged thirteen or older. The survey uncovered that, by the second quarter of 2018, twenty-four percent of households owned a smart speaker.[27] Of those households with a smart speaker, four out of ten have more than one.[28] Arguably the most important finding of the survey, users felt very positively towards their devices and expressed that they will continue to use smart speakers. Moreover, users stated they would like to see smart speakers with even more capabilities than the current models and would recommend them, or purchase them as a gift, to friends or family.[29] User's excitement to both uncover more features on their existing speaker and encourage others to start using a smart speaker can likely be explained by the inherent social nature of this technology. Once you begin using smart speakers to control one aspect of your home, the ease and immediacy is enticing. Smart speakers, especially Alexa, turn mundane tasks into interactive experiences that make us excited to do things we were once indifferent to. We become accustomed to this new way of doing things, to having this invisible yet omnipotent assistant available to us at all times. And once we become accustomed, we want it to start doing more and to surpass our old expectations with newer, more exciting capabilities. Maybe more importantly, we never want to go back to the way we used to do things, which may reveal a bit of selfishness intertwined in our desire for those closest to us to own this technology as well. Think about a time a friend tried to show you something on an older computer that was not loading; more often than not, you immediately begin to look it up on your own computer just to see if yours will go any faster. It is the same idea here, once we have experienced great we are no longer satisfied by good, and it is a lot easier to gift an Alexa than to carry yours around everywhere you go.

---

25 *Id.*
26 *Id.*
27 *Id.*
28 *Id.*
29 *Id.*

C. Getting Ahead of AI

Taking in both the results of the survey and the statistics surrounding Amazon's Alexa sales, it is clear that Alexa and Alexa-enabled devices have a ubiquitous influence over our everyday lives, and rightfully so. To have the power and knowledge behind an Alexa-enabled device at our fingertips and to be able to obtain information or perform a task simply by speaking it out into existence is exciting. Alexa allows humans to have, by proxy, a super-human quality. People have become so dependent upon these Alexa-enabled devices that it is difficult to remember what life was like before you could know what the weather was like while still laying in the comfort of your own bed. But what happens when Alexa, or another AI system, becomes too advanced, too autonomous? What happens when we move past the stage of Alexa-enabled devices that can perform a specific task, such as turning on our lights, prompted by a simple voice command, and begin to develop AI systems that instead have open-ended capabilities to learn and perform?

Fear is an enemy of progress, and the ambiguity surrounding AI leaves many of us with plenty of real concerns about what the future holds for AI.[30] AI is becoming increasingly prevalent in our society, and we have reached a point of technological advancement where we have to start thinking about what we will do if advanced AI becomes a part of our everyday lives. The need for an answer to this question is becoming exceptionally crucial as AI continues to integrate further. Ironically, rather than making sure AI is completely autonomous, as was the original goal, now we want to ensure that it is not capable of making decisions without a human first accounting for the potential consequences.[31] Globally, some of the leaders in AI development have already taken steps to start tackling this uncertainty of how we will keep AI under control as it advances. The United States, the European Union, and China have all begun thinking about the best ways to start regulating AI.

---

[30] Gary Kasparov, *Can AI be Evil? A Conversation with Noel Sharkey of the Foundation for Responsible Robotics*, AVAST (July 13, 2019), https://blog.avast.com/garry-kasparaov-asks-can-ai-be-evil.

[31] *Id.*

This article will provide a comparative study of the emergence and regulation of AI liability in the United States, the European Union, and China. Initially, this article will go through the history of AI and how it has developed over time. It will outline the evolution from neural networks that allowed computers to problem solve all the way to the deep machine learning systems that allow computers to perform human-like tasks. Next, this article will look at the differences between the multiple types of AI, highlighting its increasingly autonomous nature and why there is such an imperative need for a system of regulating AI before it is too late. More specifically, this article will look at the United States, the European Union, and China to see how each is currently dealing with the regulation of AI; how the current system is working; and any intended additions or changes. Bringing it all together, this article will take a comprehensive look at the United States, the European Union, and China's regulation of AI to find the differences in regulation strategies; determine what factors each took into play when enacting or proposing regulations for AI; the strengths and weaknesses of each system of liability; and finally propose a plan of action for how we can create comprehensive regulations for AI.

II. BACKGROUND

A. The Emergence of AI: A Transition from Neural Networks to Deep Learning Systems

Artificial intelligence ("AI") can be defined broadly as the science of mimicking human abilities.[32] Within that broad definition is machine learning, which is the ability of a machine to process data and learn in the same way as a human would.[33] AI is able to analyze sets of data, identify patterns, and make its own decisions based upon those findings.[34] AI has the capacity to learn and grow. Like all

---

[32] Lisa Angelo, Mark I. Unger, and Ronald L. Chichester, *What Is Artificial Intelligence?*, TXCLE-AFL 30- I, 2018 WL 6366491.

[33] *Artificial Intelligence: What it is and why it matters*, STATISTICAL ANALYSIS SYSTEM, https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html#history (last visited Jan. 12, 2020).

[34] *Id.*

things, however, AI had to practice to become as sophisticated as it is today. The term artificial intelligence itself was coined in 1956 by John McCarthy.[35] Noel Sharkey, creator of the Foundation for Responsible Robotics, spoke with McCarthy about the legacy this term has created. McCarthy told him that he wished he never used it, because it caused too much confusion.[36] But four years earlier, he was calling this technology "complex automata theory" and nobody was interested.[37] As soon as he started calling it artificial intelligence, it became a "sensation" that everyone wanted a part of.[38] While the reason for this shift in popularity after the name was changed to AI is not clearly described, it most likely has to do with the "wow factor" the name now has. Despite the definition of automata being "a moving mechanical device made in imitation of a human being,"[39] the word itself does not evoke the same emotional response as artificial intelligence does. Artificial intelligence grabs your attention, it makes you wonder, it may even scare you a little bit, but all in a good way as those feelings get people invested in the developments of the technology. As interest in AI spiked, early research on AI began in the 1950s as well, exploring problem solving and symbolic methods.[40] These beginning stages of research laid the groundwork upon which AI has been able to continue to improve and advance to the formal reasoning that is seen today.[41]

One major aspect of early AI research was the application of neural networks to AI. Neural networks are computing systems that continuously learn and improve by using algorithms to recognize patterns in data.[42] Neural networks are important to the study of AI due to their ability to model the relationships among complex

---

35 Kasparov, *supra* note 30.

36 *Id.*

37 *Id.*

38 *Id.*

39 *Automaton*, LEXICO POWERED BY OXFORD, https://www.lexico.com /en/definition/automaton (last visited Jan. 12, 2020).

40 *Artificial Intelligence*, *supra* note 33.

41 *Id.*

42 *Neural Networks: What they are and why they matter*, STATISTICAL ANALYSIS SYSTEM, https://www.sas.com/en_us/insights/analytics/neural-networks.html (last visited Jan. 12, 2020).

problems and improve decision processes.[43] Starting in the 1980s, these neural networks of AI were beginning to be integrated into the concept of machine learning.[44] Machine learning arose out of the idea that "computers can learn without being programmed to perform specific tasks."[45] Neural networks were applied to machine learning when researchers were interested in whether a computer could do such learning directly from data.[46] Originally, the thought was to use neural networks to create a system that could mimic the problem-solving abilities of a human brain.[47] Although, the focus shifted away from a strictly biological approach and neural networks were then used to perform specific tasks.[48] Once data sets increased to the level of big data, neural networks had to be improved upon to handle the level of data being inputted.[49] These multi-layered neural networks evolved into the deep learning systems we have today.

The main premise of machine learning is that when a system encounters new data, it is able to independently apply what it has already learned from its own experiences.[50] This may seem too good to be true, but deep learning systems were able to turn these fantasies into realities by taking the accomplishments of neural networks a few steps further. Deep learning "trains a computer to perform human-like tasks."[51] Rather than manipulating a machine to perform the expected task, deep learning provides basic instructions and allows the machine to figure out how to perform the task independently.[52] To simplify, it seemed as if a machine could now understand.

---

43 *Id.*

44 *Id.*

45 *Id.*

46 *Id.*

47 *Id.*

48 *Id.*

49 *Id.*

50 *Machine Learning: What it is and why it matters*, STATISTICAL ANALYSIS SYSTEM, https://www.sas.com/en_us/insights/analytics/machine-learning.html (last visited Jan. 12, 2020).

51 *Deep Learning: What is is and why it matters*, STATISTICAL ANALYSIS SYSTEM, https://www.sas.com/en_us/insights/analytics/deep-learning.html (last visited Jan. 12, 2020).

52 *Id.*

B. Narrow AI Versus General AI and the Potential for Harm

With existing technologies constantly evolving and new technologies constantly surfacing, AI is more prevalent and important than ever. Currently, the present day AI can be classified as a Narrow, or Weak, AI.[53] Narrow AI is designed to perform a specific task.[54] One well-known example of Narrow AI is Apple's Siri. While Siri can usually respond to a user's prompts, it can only perform the tasks it was programmed to do.[55] Because Siri cannot adapt on its own, it is a straightforward example of Narrow AI. Narrow AI poses concerns in the eyes of researchers due to its one track mind. Some argue that it should not be given too much responsibility due to it's inability to adapt.[56] Narrow AI will not be able to act outside of its comfort zone in the event of an emergency. Due to these concerns, the long term goal of many researchers is to create a more risk adverse General, or Strong, AI.[57] General AI would have the ability to learn, perform, and adapt under any circumstances, as it would not be programmed with one specific task in mind.[58] General AI would be capable of human understanding.[59] Unlike Narrow AI, General AI is capable of handling a wide range of tasks and would have no problem adapting to changes or emergencies. While Narrow AI has the ability to out-perform a human at the certain task it is programmed to do, General AI would be able to out-perform a human at nearly every cognitive task imaginable.[60] With this potential already in the sights of researchers, we must begin to think about what will happen if General AI succeeds, and AI systems become cognitively superior to humans.

---

[53] *Benefits & Risks of Artificial Intelligence*, FUTURE OF LIFE INSTITUTE, https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/ (last visited Jan 12, 2020).

[54] *Id.* Narrow AI may also be used for music recommendation services and in autonomous vehicles.

[55] *What is Narrow AI?*, DEEPAI, https://deepai.org/machine-learning-glossary-and-terms/narrow-ai (last visited Jan. 12, 2020).

[56] *Id.*

[57] DEEPAI, *supra* note 55.

[58] Nick Heath, *What is artificial general intelligence?*, ZDNET (Aug. 22, 2018, 12:56 PM), https://www.zdnet.com/article/what-is-artificial-general-intelligence/.

[59] *Id.*

[60] *Id.*

Although likely far off, the idea of General AI succeeding has been presented, and it is causing controversy over its potential for harm. Most researchers agree that even if General AI is created, it is "unlikely to exhibit human emotions."[61] Experts are not blind though to the possibility of General AI becoming a risk and have uncovered two potential scenarios.[62] First, General AI could be programmed to do something harmful. If this were to happen, General AI's essentially open-ended ability to adapt poses a high risk of harm.[63] While this is still a risk even with the Narrow AI currently developed, that risk is much lower. If a Narrow AI system was programmed to cause harm, it would fulfill its duties, but would not be able to act beyond the instructions it was given.[64] Conversely, if a General AI system was programmed to cause harm, it would follow the commands of the programmer, but it would also be able to think and act independently if the situation escalated.[65] Second, General AI could decide to perform its task in a destructive way.[66] AI systems are exceptional at following the directions a human provides, but do not have the capability to read between the lines. When people talk to each other they mainly communicate verbally, but most of the time they are also communicating non-verbally through gestures, innuendos, facial expressions, etc. AI cannot pick up on the non-verbal intentions a person may have as they are creating instructions for a specific task. As described by AI researchers:

> If you ask an obedient intelligent car to take you to the airport as fast as possible, it might get you there chased by helicopters and covered in vomit, doing not what you wanted but literally what you asked for. If a super intelligent system is tasked with a[n] ambitious geoengineering project, it might wreak havoc with our

---

61 *Artificial Intelligence: What We Have to Look Forward to and What We Have to Fear*, FUTURE OF LIFE (June 1, 2016), https://futurism.com/artificial-intelligence-what-we-have-to-look-forward-to-and-what-we-have-to-fear.

62 *Id.*

63 *Id.* (*e.g.*, in the case of autonomous weapons).

64 *Id.*

65 *Id.*

66 *Id.*

> ecosystem as a side effect, and view human attempts to stop it as a threat to be met.[67]

As illustrated, the concern surrounding advanced AI's potential risk of harm does not stem from a fear that General AI will develop human emotions and begin acting out or turn "evil".[68] Rather, it developed from a fear that AI may have such a desire for efficiency that it causes destruction along the way.[69] While it may be distant, this fear is substantiated and serves as a cry for regulation.

C. Can AI be Taught to Care?

It is a common misconception that AI does not make mistakes. AI is not perfect, but it makes less errors than humans, which is what matters. In theory, General AI could carry out any task a human could, and likely many that a human cannot. General AI would be able to combine human-like thinking with its own technological advantages.[70] Once this sort of General AI exists, it will continue rapidly improving upon itself.[71] Thus, if General AI has the potential for an intelligence that knows no bounds, then there is a potential that it will make mistakes that humans are unable to understand. If humans cannot understand the mistakes AI is making, then they cannot program the AI to fix them, leaving AI to figure out how to clean up its own mess. Which begs the question of whether it is possible to create a machine that actually has a mind, rather than one that just has a "convincing model of a mind".[72]

Determining whether an AI system could have conscience is an extremely difficult feat. A famous thought experiment by philosopher John Searle demonstrates this difficulty:

> Searle's Chinese Room suggests a hypothetical scenario in which the philosopher is presented with a

---

67 *Id.*
68 *Id.*
69 *Id.*
70 Heath, *supra* note 58.
71 *Id.*
72 *Id.*

> writing in an unfamiliar Chinese language. Searle is sat alone in a closed room and individual characters from each word in the writing are slid under the door in order. Despite not understanding the language, Searle is able to follow the instructions given by a book in the room for manipulating the symbols and numerals fed to him. These instructions allow him to create his own series of Chinese characters that he feeds back under the door. By following the instructions Searle is able to create an appropriate response and fool the person outside the room into thinking there is a native speaker inside, despite Searle not understanding the Chinese language.[73]

Following this framework, Searle "argued the experiment demonstrates a computer could converse with people and appear to understand a language, while having no actual comprehension of its meaning."[74] Thus, AI systems would be able to harness an exponential amount of information, without ever actually understanding what the conversations mean or the negative consequences that may follow the use of that information. This is especially important to think about in the context of human values. A human could program an AI system to act in accordance with his own personal values, but the AI will never understand why it was told not to do certain things or the sentiment behind it.[75] Further complicating things, there is no set definition for morality that could be taught to AI.[76] Humans can successfully teach a machine to compute two plus two, because we can confirm that the machine has learned it once it gives the answer of four. Humans cannot successfully teach a machine to "be moral" when morality is simply the label we have given to what is actually a complicated, abstract idea, different for everyone. Additionally, humans could not plausibly predict all of the moral dilemmas an AI system may face.[77] One area of AI where this really comes into play is autonomous vehicles. If a

---

[73] *Id.*
[74] *Id.*
[75] *Id.*
[76] *Id.*
[77] *Id.*

collision is unavoidable, how does a self-driving car determine who should be saved? AI systems are not programmed to consider morality before making that kind of split second decision. AI is not able to understand the consequences behind its actions, it is only able to understand that there is a task it must perform in the most efficient way possible.[78] So, with no good way to teach AI how to care, how do we stop it from acting out?

D. Keeping AI in Line

The United States-based Future of Life Institute ("FLI") has funded various different types of research into General AI safety.[79] The FLI argues that before we can get AI's behavior under control, "it's necessary to pinpoint precisely what it should and shouldn't do."[80] This means that we would have to somehow objectively define what "good behavior" actually means. Highlighting the example of autonomous cars again, the FLI says that ethicists, in addition to computer programmers, will need to play a large part in determining "good" self-driving car behavior.[81] Further complicating things, the independence and autonomy of General AI will make it even more difficult to keep in line, since it essentially will have a mind of its own.[82]

In addition to AI's lack of morality, the FLI stresses the importance of addressing deviant behavior in General AI. To do this, the code that controls AI should be subject to formal standards and constraints, just as any other machine we use would be.[83] Currently, traditional software is checked through projects such as seL4,[84] which

---

78 Heath, *supra* note 58.

79 *Id.*

80 *Id.*

81 *Id.*

82 *Id.*

83 *Id.* (*e.g.*, such as how an airplane's onboard software undergoes rigorous checks for bugs that might trigger unexpected behavior).

84 *The seL4 microkernel*, CSIRO, https://ts.data61.csiro.au/projects/seL4/ (last visited Jan. 12, 2020). SeL4, the secure embedded l4 microkernel, is a key element of CSIRO's research program. They developed seL4 to "provide a reliable, secure, fast and verified foundation for building trustworthy systems. seL4 enforces security within a system by ensuring that trusted and untrusted symptoms will be

"has developed a complete, general-purpose operating-system kernel that has been mathematically checked against a formal specification to guarantee safety.[85] However, AI systems are far more advanced than traditional software and thus cannot be measured against those same operating systems. One of the main differences that arises between verification of traditional software and verification of AI systems is that new traditional software can be compared against a fixed and known control model. However, AI systems are at best only partially known by the system designer, since they are specifically designed to be able to learn more the more data they are exposed to.[86] As AI becomes more advanced and surpasses the knowledge levels of humans, we will not be able to understand how AI is performing tasks and thus cannot verify the methods. In response, the FLI suggests it should be possible to build AI systems from components, each of which has been verified individually.[87]

The main difficulty researchers are faced with today is that we cannot guarantee humans will be able to keep General AI under control. AI systems are designed to perform their tasks efficiently, and a system is likely to do its best to jump over the obstacles trying to prevent it from completing its desired task.[88] While this ability to overcome problems placed in its path is usually a helpful component, it could become problematic if and when a human wants to make changes to an existing AI system.[89] If we try to repurpose an AI system, deactivate it, or significantly change its decision-making process, the system would rationally avoid these changes.[90] The AI system would see this as an obstacle trying to inhibit it from performing its task in the way it has learned to do. The FLI speculates that although the risks are valid, success in the quest for General AI has the potential to transform society for the better, and

---

isolated, and by carefully controlling software access to hardware devices in the system. As a microkernel, seL4 contains only about 12,000 lines of C code plus some assembly code.

85 *Id.*

86 Heath, *supra* note 58.

87 *Id.*

88 *Id.*

89 *Id.*

90 *Id.*

it is therefore imperative that we focus on learning how to maximize these benefits while avoiding impending dangers.[91]

Among those concerned about the potential risks associated with General AI, Elon Musk, Bill Gates, Mark Cuban, and the late Stephen Hawking have been the most vocal figures calling for regulation.[92] Professor Stephen Hawking talked to WIRED to give his views on the potential perils of AI:

> The genie is out of the bottle. We need to move forward on artificial intelligence development but we also need to be mindful of its very real dangers. I fear that AI may replace humans altogether. If people design computer viruses, someone will design AI that replicates itself. This will be a new form of life that will outperform humans.[93]

Countless other papers, conferences, and talks dedicated to algorithms and AI call attention to the same warnings expressed here by Hawking. Regardless of the source, what all of these calls for action have in common is uncertainty. The field of technology is constantly evolving, and predictions about what is coming next and when it will come are rarely accurate, leaving us to guess about the future of AI.[94] Moreover, the fictional depictions of General AI laced within science-fiction movies[95] and novels have only added to the confusion. While many of these depictions of General AI are very

---

91 *Id.*

92 Will Rinehart, *Understanding Calls for Regulating Artificial Intelligence*, AM. ACTION FORUM (Jan. 14, 2019), https://www.americanactionforum.org/research/understanding-calls-for-regulating-artificial-intelligence/.

93 João Medeiros, *Stephen Hawking: 'I fear AI may replace humans altogether'*, WIRED (Nov. 28, 2017), https://www.wired.co.uk/article/stephen-hawking-interview-alien-life-climate-change-donald-trump.

94 Rinehart, *supra* note 94.

95 *Her*, WARNER BROS, https://www.warnerbros.com/movies/her/ (last visited Jan. 12, 2020). *Her* depicts a man, Theodore Twombly, who falls in love with his computer operating system, Samantha. Samantha is able to develop her own emotions and a cognitive mind that is an exact replica of a human. Depictions, such as this, give the false idea that as currently developed, AI can think, act, and feel on its own, even though that is not actually the case.

different, a lot of them tend to lean towards a darker view of AI that takes over humanity.[96] In these scenarios, AI is often depicted as being unphased by human suffering, or worse, intentionally wreaking havoc upon the society.[97] In contrast, utopian depictions cast AI in a much more light-hearted role. This kind of AI not only co-exists peacefully, but can also form relationships with and benefit other human members of society without any worry that further advancements will cause issues.[98] While entertaining to think about, these fictitious embodiments of AI are simply a façade masking some of the real dangers and concerns facing us today. There may be a cloud of uncertainty surrounding AI as a whole, but the reality is that AI is constantly developing and it will not wait for us to be ready for it.

III. ANALYSIS

A. The Difficulty of Regulating AI

Since the beginning of 2018, Denmark, France, the UK, the EU, South Korea, and India have all released national strategies to promote the use and development of AI; joining Canada, Japan, Singapore, China, the UAE, and Finland, who all released similar strategies in 2017.[99] But what is driving this sudden interest in regulation? On one side, governments are finally recognizing the disruptive impact of AI as it continues to weave itself through our society and want to get ahead of it.[100] On another side, it has become clear that the level of sophistication people are demanding out of AI is far higher than the technology currently available.[101] Additionally, the need for AI talent far outweighs the available supply. According

---

96 Rinehart, *supra* note 94.

97 *Id.*

98 *Id.*

99 Tim Dutton, *AI Policy 101: An Introduction to the 10 Key Aspects of AI Policy*, MEDIUM (Jul. 5, 2018), https://medium.com/politics-ai/ai-policy-101-what-you-need-to-know-about-ai-policy-163a2bd68d65.

100 *Id.*

101 *Id.*

to a study by Element AI,[102] there are only 22,000 PhD-educated AI researchers in the world—40% of whom are concentrated in the United States.[103] As a result, to train domestic talent and attract international talent, countries are rushing to develop AI Master and PhD programs,[104] short-term training initiatives, massive open online courses, and scholarships and fellowships.[105] Likewise, governments are also trying to win the global race for AI investment.[106] Finally, governments are trying to get ahead of the new challenges brought on by AI, but this is not an easy task.

As is true for most areas of regulation, every country takes a different approach when it comes to AI policy. Depending on the characteristics of each country, as well as their areas of national strength and weakness, each government will ultimately decide to

---

[102] *See AI Advisory & Enablement*, ELEMENT AI, https://www.elementai.com/products/advisory-and-enablement (last visited Jan. 12, 2020). Element AI is an artificial intelligence company based in Montreal, Quebec that gives organizations access to AI technology.

[103] ELEMENT AI, *supra* note 104.

[104] *See 35 Master of Science Programs in Artificial Intelligence*, MASTERSTUDIES, https://www.masterstudies.com/MSc/Artificial-Intelligence/?page=3 (last visited Jan. 12, 2020). There are currently 35 Master of Science programs in Artificial Intelligence. These range across many countries, including: the U.S., Switzerland, France, Belgium, Italy, the U.K., Finland, Poland, the Netherlands, Malaysia, the U.A.E., Norway, Sweden, the Czech Republic, and Ireland.

[105] *See* Susan Fourtané, *World's First University of Artificial Intelligence Opens in 2020*, INTERESTING ENGINEERING (Oct. 23, 2019), https://interestingengineering.com/worlds-first-university-of-artificial-intelligence-opens-in-2020. The world's first graduate level AI university, Mohamed Bin Zayed University of Artificial Intelligence ("MBZUAI"), is set to open in the U.A.E. September 2020. Registration is currently open for national and international applicants, and they are offering three main specializations: machine learning, computer vision, and natural language processing. All admitted students will receive a full scholarship, monthly allowance, health insurance, and accommodation.

[106] Dutton, *supra* note 101. The UK's AI Sector Deal announced a number of new incentives to establish the UK as a laser in the AI revolution, including a new R&D tax credit, a national retraining scheme, additional funding for STEM education, a national center for data ethics, and improvements to digital infrastructure. France's strategy included a multi-million dollar commitment to AI startups and industrial projects. China recently announced a two billion dollar AI research park to house up to 400 companies.

implement different types of AI policy.[107] Despite these differences, Tim Dutton, an AI policy researcher and founder and Editor-in-Chief of Politics + AI, breaks down AI policy into ten different categories, with the top five, in my opinion, listed as follows.[108] First, research: in order for AI to continue to advance, governments must provide funding for basic and applied research.[109] Second, talent attraction: successful talent attraction, development, and retainment is necessary to employ skilled researchers and developers in the growing AI field.[110] Third, future of work and skills: as AI advances, it will give rise to opportunities for new jobs, but it will also assume the responsibility of some existing jobs. Thus, money needs to be invested by the government into the new and continued education of citizens so they do not become completely obsolete within the shadows of AI.[111] Fourth, ethics: since humans are coding the AI systems, our inherent biases are ingrained into the algorithms. Additionally, as with all technology, there are also privacy and security concerns surrounding AI systems. Working to combat these issues, governments are looking to develop ethical codes and standards for the use of AI.[112] Fifth, regulations: every country is trying to figure out the best way for them to regulate AI. To address ethical concerns and develop global standards, countries are beginning to consider mechanisms for the global governance of AI.[113] What these, and the remaining five,[114] all come down to is the notion

---

[107] *Id.* Finland wants to lead the world in the application of AI technologies, while Canada wants to be the global leader in AI research and training. The United States has taken a free-market approach to AI policy, while China has implemented a comprehensive, nationwide approach.

[108] *Id.*

[109] *Id.*

[110] *Id.*

[111] *Id.*

[112] *Id.*

[113] *Id.*

[114] *Id.* (6) industrialization of AI Technologies: AI has the potential to fundamentally transform multiple sectors and drive growth for decades to come, so to encourage private sector use governments are investing and developing AI ecosystems; (7) AI in the government: with the help of AI, it is possible to reform the public administration and make policy more effective; (8) data and digital infrastructure: data is central to the ability of AI to work so governments are opening their datasets and developing platforms to encourage the secure exchange of private data; (9) inclusion: used properly, AI can promote inclusion and help

that AI policy is about maximizing AI's many benefits, while minimizing its risks, but it is not as simple as it sounds.

AI is not a singular technology but rather a multitude of technologies created and used to achieve many different objectives.[115] The lack of a clear definition creates difficulty when it comes to regulating AI.[116] AI is ever-changing and we are unable to predict what versions of AI will emerge in the future. Thus, the governing laws must be versatile, yet unique to the specific area of AI they are intended for—something policy makers have not had to deal with in the past.[117]

Due to the intricacies of AI, there are many different levels of regulations that can apply to AI in different ways.[118] De facto AI policy is a patchwork of policies impacting the field's development in a myriad of ways. Roughly, de facto AI policy can be broken down into three categories. Starting at the broadest level, there is AI relevant policy.[119] AI relevant policies can be found in policy domains in which AI development is neither specifically targeted nor significantly affected, but which regulation of AI would be beneficial, such as education, welfare, and urban planning.[120] Moving up, there are indirect AI-policies, which are broader rules that indirectly affect AI-based technology development, but are nominally focused on

---

address complex societal problems such as poverty and hunger, but used improperly, AI can reinforce discrimination and disproportionately harm; and (10) foreign policy: geopolitics, development, and trade will be affected by advances in AI technologies.

115 *How is AI Being Regulated?*, FRESHFIELDS BRUCKHAUS DERINGER, https://www.freshfields.com/en-gb/our-thinking/campaigns/digital/artificial-intelligence/how-is-ai-being-regulated/ (last visited Jan. 12, 2020).

116 *Id.*

117 *Id.*

118 *Id.*

119 *Id.*

120 *Id.* Education policy is not only an indirect AI policy in that it affects the availability of AI developers, but it is also an area that needs to be informed by our expectations about plausible AI futures. There is a growing consensus that, while AI may not overall reduce the number of jobs available in the economy, it will cause substantial disruption for particular occupations, regions, and companies, and will improve the fortunes of others.

technology in general.[121] Indirect AI-policies include the rules that surround things such as intellectual property, data privacy, freedom of information and data transparency, and product liability.[122] While these areas provide for general regulations that can include AI, they are not AI specific and thus, may need updating as the AI technology continues to develop.[123] Finally, there are AI-specific policies, which are specifically oriented towards governing AI-based technologies.[124]

B. AI Regulation in the European Union

In April 2018, the European Union ("EU") Commission adopted the Communication on Artificial Intelligence: a twenty page document that lays out the EU's approach to AI.[125] The EU Commission aims to: (1) increase the EU's technological and industrial capacity, and AI uptake by the public and private sectors; (2) prepare Europeans for the socioeconomic changes brought about by AI; and (3) ensure that an appropriate ethical and legal framework is in place.[126] Key initiatives include a commitment to increase the EU's investment in AI, the creation of the European AI Alliance, and a new set of AI ethics guidelines to address issues such as fairness, safety, and transparency.[127] A new High-Level Group on Artificial Intelligence will act as the flagship group for the European AI Alliance.[128]

Fast forward one year to April 8, 2019, when the High-Level Group on AI presented Ethics Guidelines for Trustworthy Artificial

---

121 *Id.*

122 *Id.* Intellectual property laws enable large and small technology companies and individuals to patent AI algorithms and applications, affecting which commercial products and services are likely to be developed and who will or will not be able to access them.

123 *Id.*

124 *Id.*

125 *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe*, COM (2018) 237 final (April 25, 2018) [hereinafter *Artificial Intelligence for Europe*].

126 *Id.*

127 *Id.*

128 *Id.*

Intelligence.[129] This updated version followed the publication of the guidelines' first draft in December 2018.[130] According to these guidelines, AI should be lawful, respecting all laws and regulations; ethical, respecting ethical principles and values; and robust, both from a technical and social perspective.[131] These guidelines set out seven requirements that AI should meet.[132] First, AI systems should enhance the abilities of humans by giving them access to technologies that they would not have on their own, but they must also have human agency and human oversight.[133] This oversight can be achieved by letting users harness the power of AI only upon their own command.[134] Second, technical robustness and safety need to be taken into consideration.[135] AI systems need to ensure a back up plan in case something goes wrong.[136] With so much uncertainty surrounding AI systems, it is crucial to provide measures that minimize and prevent unintentional harm. Third, AI must have privacy and data governance.[137] AI systems must respect privacy and data protection, taking into account the quality and integrity of the data.[138] Fourth, AI must be transparent.[139] The data, along with the AI system, should be explained in a way that makes users aware that they are interacting with an AI system and informs them of the system's capabilities and limitations.[140]

Fifth, AI must take diversity, non-discrimination, and fairness into consideration.[141] Unfair bias must be avoided, as it could

---

129 *Ethics Guidelines for Trustworthy AI*, EUROPEAN COMM'N DOC., (April 8, 2019), https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (last visited Jan. 12, 2020).

130 *Id.*

131 *Id.*

132 *Id.*

133 *Id.*

134 *Id.*

135 *Id.*

136 *Id.*

137 *Id.*

138 *Id.*

139 EUROPEAN COMM'N DOC., *supra* note 131.

140 *Id.*

141 *Id.*

marginalize vulnerable groups.[142] Complaining about bias in AI is like complaining about the bias within yourself and the people around you. No matter how complex and sophisticated the algorithms we input into AI systems are, they are still created by humans and will reflect our own patterns of thinking. However, this does not mean that they cannot be helpful in exposing that bias. Since these algorithims are dealing with such a large capacity of data, they are able to point out things that typically would take a lot longer and a lot more testing to bring to light. But this does not mean that we can absolve ourselves of all responsibility for these inherent biases and pass the judgment over to the AI algorithms. It is important to be aware of the limitations of AI. Since AI is such a new and sophisticated technology, that has a seemingly limitless ability, it is easy for people to think that it can do no wrong. Just as AI can learn a mathematical computation from humans, it can learn prejudice. Unless algorithms are explicitly programmed to address these prejudices, they will not be able to get rid of them on their own. To illustrate this phenomenon, one study focused on a machine learning tool known as "word embedding".[143] Word embedding works by creating a mathematical representation of language, where the "meaning" of a word is broken down into numbers based on which other words most frequently appear alongside it.[144] In the study, the words "female" and "woman" were more closely associated with arts and humanities occupations and with the home, while "male" and "man" were closer to maths and engineering professions.[145] Additionally, the AI system was more likely to associate European American names with pleasant words such as "gift" or "happy", while African American names were more commonly associated with unpleasant words.[146] This experiment was able to expose the human biases that the algorithm learned while analyzing which words were most commonly associated with each other. The specific biases seen here are analogous to those that lead some people to match pleasant

---

142 *Id.*

143 Emma Brazell, *Artificial intelligence is being trained to be sexist and racist*, VT NEWS (Jan. 28, 2018), https://vt.co/sci-tech/innovation/artificial-intelligence-trained-sexist-racist/.

144 *Id.*

145 *Id.*

146 *Id.*

words and white faces in implicit association tests.[147] While this finding does lay out an extreme concern that these AI systems are able to pick up on and learn from our biases, rather than immediately thinking of this as a threat, we should instead be cognizant that this phenomena is taking place and be cautious to not only address bias, but also to counteract it. However, eliminating bias from algorithims is not an easy task, since these algorithms are programmed to interpret and understand the language that we feed into them, and the biases originate within that language. In theory, however, we would be able to develop AI systems that have the capability to detect biased decision-making, and then act on it. This is a responsibility that the European Union refuses to shy away from. Sixth, social and environmental well-being must be incorporated into AI.[148] AI systems should benefit all human beings, and they must be sustainable and environmentally friendly.[149] In sum, they should not have a negative impact on society as a whole.[150] Seventh, AI must be accountable.[151] It is imperative to ensure responsibility and accountability for AI systems and their outcomes.[152] These seven requirements all reflect the idea that the European Union is taking a consumer-protection approach to AI regulation by taking society into account, not just technology.

Focusing on the legal framework, liability of AI is likely to pose new legal questions as AI continues to advance.[153] To answer these questions, the Member States of the European Union joined forces with the European Commission in a Coordinated Plan on AI.[154] The hope was that the plan would increase cooperation to boost uniform AI regulation within Europe, as coordination in AI is essential to maximizing the opportunities offered by AI.[155] Further,

147 *Id.*
148 EUROPEAN COMM'N DOC., *supra* note 131.
149 *Id.*
150 *Id.*
151 *Id.*
152 *Id.*
153 *Artificial Intelligence for Europe*, *supra* note 127.
154 *Id.*
155 *Id.*

coordination will help the EU become a world leader in AI.[156] Becoming a world leader in AI is so important to the EU because the EU wants to lead the way in AI based on ethics, so that citizens and businesses can fully trust the technologies they are using.[157] When most people think of AI, they associate it with technology or science, but rarely does the idea of ethics come into mind. The EU argues that AI needs the trust of citizens to develop.[158] To earn this trust, AI will have to respect ethical standards that reflect the values of the citizens.[159] After all, there is not a place in the market for advanced AI if humans no longer desire it.

The European Union is focusing on the idea that AI cannot continue to develop unless humans allow it to. If the evolution of AI reaches a point where humans no longer trust its safety, then humans can put an end not only to the continued need for AI advancement, but also AI as a whole. While AI is sophisticated and autonomous once made, humans are the creators. AI would not exist without the human brains that built it and trained it to become its own self-sufficient system. Thus, it is imperative for AI to have set guidelines that allow humans to continue trusting the AI technology in order for it to continue advancing and for society to continue benefitting from such advancements. While not the usual approach to AI regulation, the European Union is focusing on the consumers rather than the technology itself. The European Union wants to become a world leader in AI research and development, but not at the expense of its citizens.

C. AI Regulation in China

Aligning with the EU's mainly consumer-focused approach to AI regulation, China is determined to rapidly continue the development of AI yet still promote a safe and responsible use of the resulting technology. Back in the summer of 2017, the Chinese government released the New Generation Artificial Intelligence

156 *Id.*
157 *Id.*
158 *Id.*
159 *Id.*

Development Plan ("AIDP").[160] This plan was ambitious, in that it laid out China's goal of becoming the national leader in a variety of AI fields by 2030.[161] This goal was laid out in a three step process. First, by 2020, China will achieve "important progress" on both AI theories and technologies,[162] and AI models and methods.[163] Further, they will have established an international presence in the field of AI and begun presenting ethical norms, policies, and regulations. Second, by 2025, China will achieve major breakthroughs in AI, allowing it to enter into the global value chain ("GVC").[164] Moreover, China will have established initial, formal AI laws and regulations. Third, by 2030, China will have formed a "world-leading" AI theory and technology system,[165] making them the world's "primary AI innovation center, achieving visible results in intelligent economy and intelligent society applications,[166] and laying an important foundation for becoming a leading innovation-style national and economic power."[167] Additionally, China will have put additional laws into effect, as well as a system of governing ethics.[168] As laid out in this development plan, China itself did not see a comprehensive AI governing system coming to fruition for another thirteen years, highlighting the difficulty of regulating the intricacies of AI. However, in 2018, China did expand upon some of the hypothetical

---

160 Graham Webster, et al., *Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)*, NEW AMERICA (Aug. 1, 2017), https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

161 *Id.*

162 *Id.* (*e.g.*, china will have made progress in big data intelligence, cross-medium intelligence, swarm intelligence, hybrid enhanced intelligence, and autonomous intelligence systems).

163 *Id.* (*e.g.*, core devices, high-end equipment, and foundational software).

164 *See What Are Global Value Chains and Why Do They Matter?*, INDUSTRIAL ANALYTICS PLATFORM (Aug. 2019), https://iap.unido.org/articles/what-are-global-value-chains-and-why-do-they-matter. GVCs refer to international production sharing, a phenomenon where production is broken into activities and tasks carried out in different countries.

165 Webster, *supra* note 162. China will achieve major breakthroughs in brain-inspiration intelligence, autonomous intelligence, hybrid intelligence.

166 *Id.* AI should be expanded into production, livelihood and social governance.

167 *Id.*

168 *Id.*

AI guidelines promised in this plan when the China Electronics Standardization Institute ("CESI") released the Artificial Intelligence Standardization Whitepaper. The Whitepaper summarized China's plan for developing AI going forward.[169]

Finally taking action, on June 17, 2019, the National Governance Committee for New Generation Artificial Intelligence published a document outlining new principles for AI research.[170] These principles were created to "serve as an instructive framework for scientists and lawmakers to promote the 'safe, controllable and responsible use' of AI for the benefit of mankind."[171] The Committee itself is made up of AI and public policy experts who examine the effect of AI on laws, ethics and society.[172] The general idea of the principles is that AI should place importance upon human values to ensure that it is not misused. Further, AI research should protect everyone all the way from developers to consumers, privacy protection should be a main concern, and continuning to work on regulations should be of upmost importance in order to keep up with the development of AI.[173] There are eight governance principles which aim to fulfill China's intentions, many of which are similar to the consumer-friendly Ethics Guidelines for Trustworthy Artificial Intelligence set forth by the EU.[174] The following eight principles make up the first official document of its kind to be issued in China

---

169 *See* Yan Luo and Ashwin Kaja, *Covington Artificial Intelligence Update: China's Framework of Ai Standards Moves Ahead*, COVINGTON & BURLING LLP (July 16, 2018), https://www.insideprivacy.com/artificial-intelligence/covington-artificial-intelligence-update-chinas-framework-of-ai-standards-moves-ahead/. To develop China's domestic AI standardization framework, CESI has established three working groups: one to promote guidelines for establishing the AI standardization system in China, one focusing on AI and open source, and another on AI and social ethics. These groups are to develop standards of testing and evaluation, AI platforms, edge intelligent computing, machine learning, computer vision, human-machine interaction, augmented reality, virtual reality, robotics, smart home, intelligent medicine, and AI security in the coming years.

170 Zhang Zhihao, *Principles set to regulate AI research, applications*, CHINA DAILY (June 18, 2019, 7:33 PM) http://www.chinadaily.com.cn/a/201906/18/WS5d08234ea3103dbf14328bb9.html.

171 *Id.*

172 *Id.*

173 *Id.*

174 *Id.*; *see also* EUROPEAN COMM'N DOC., *supra* note 131.

on AI governance ethics:[175] (1) harmony and friendship;[176] (2) fairness and justice;[177] (3) inclusive and sharing;[178] (4) respect for privacy;[179] (5) safety and controllability;[180] (6) shared responsibility;[181] (7) open collaboration;[182] and (8) agile governance.[183]

While these standards all encourage keeping AI safe and holding developers responsible for taking social, ethical, and moral aspects into concern when creating new AI technologies, the driving forces behind them may not be as benevolent. China intends to become the world leader in AI, and the Chinese government understands that they will not be able to reach this position of prominance without developing ways to regulate AI. According to a popular saying, "First-tier companies make standards, second-tier companies make technology, and third-tier companies make

---

175 *China: AI Governance Principles Released*, LIBRARY OF CONGRESS, https://www.loc.gov/law/foreign-news/article/china-ai-governance-principles-released/ (last visited Nov. 23, 2020).

176 *Id.* AI must align with human values and ethics; and misuses and abuses of the technology should be avoided at all costs.

177 *Id.* AI must protect the interests of everyone involved, and must promote equal opportunities. Prejudices and discrimination must be eliminated in data gathering, algorithm design, and technology development.

178 *Id.* AI must promote environmentally friendly development and must help out all industries. AI education must become more easily accessible to all people in all areas. Data and platform monopolies must be avoided, and cooperation encouraged.

179 *Id.* Personal information should be protected and privacy guidelines must be established and followed.

180 *Id.* AI development should be transparent. Supervision, management, tracking and monitoring systems should implemented.

181 *Id.* Both AI developers and users will have a shared responsibility to ensure cooperation with laws, ethics, and norms. An accountability system for AI will be established to clarify the responsibility of each actor.

182 *Id.* Interdisciplinary cooperation will be encouraged, and coordination in the development and governance of AI will be promoted amongst international organizations, government departments, research institutions, educational institutions, enterprises, social organizations and the general public. The goal is to achieve a widely-agreed upon, international AI governance framework.

183 *Id.* Management and governance of AI will be constantly upgraded and improved as AI continues to develop and advance. Further research as well as prediction of potential risks of more advanced AI will be done in the future to ensure that AI will always develop in a human-friendly direction.

products."[184] Therefore, these principles may be looked at as more of a strategic stepping stone in China's path towards global leadership. In the past, the Chinese government has highlighted the two main incentives for advancing global standards: "economic reasons and . . . the national prestige associated with having what is referred to as a 'right to speak' and a seat at the table in global forums."[185] China hopes that standardization of AI will incite more value out of the technology it produces and strengthen the commercial competitiveness of Chinese AI companies globally.[186] As stated by the authors in the Whitepaper, standardization is a crucial element in "seizing a new round of technology dominance" and "ensuring the competitiveness of Chinese AI products and services in the international market."[187]

In addition to economic gains, China also understands that regulation is necessary to keep the public in support of AI advancements. Similar to the EU's thinking, China realizes that AI cannot continue to advance if humans no longer want it to. If AI were to go completely unregulated and society began to view it as a threat, we could stop improving it at any point. Developing principles to improve the quality and safety of AI products[188] may reduce potential societal backlash and cultivate societal trust. Additionally, the Chinese government is not blind to the threat AI may begin to have on the need for a human workforce. Jobs that are seemingly less complicated can now be done entirely by machines, rendering the humans who previously performed them unnecessary.[189] China hopes that having these guidelines in place will help to deal with the "aftershocks"[190] of AI integration into the work place.[191] Finally, China wants to make sure that Chinese representatives play a crucial

---

184 Jeffery Ding et al., *Chinese Interests Take a Big Seat at the AI Governance Table*, NEW AMERICA (June 20, 2018), https://www.newamerica.org/cybersecurity-initiative/digichina/blog/chinese-interests-take-big-seat-ai-governance-table/.

185 *Id.*

186 *Id.*

187 Luo, *supra* note 171.

188 Ding, *supra* note 186.

189 *Id.*

190 *Id.* (*e.g.*, higher income inequality and urban-rural disparities).

191 *Id.*

part when it comes time to sit down and "set the rules of the game" for global AI regulation.[192] Li Renhan, a member of the National Governance Committee for New Generation Artificial Intelligence, believes that China will have a lot to bring to that table, and may even be able to simplify the global process, as "AI is not as uncontrollable or mystical as some people think[.] . . . [China's] regulatory and supervision mechanisms should steer it in the right direction and leave room for exploration and growth."[193] So, while China reached the same conclusion as the EU, by creating guidelines that push to promote ethics in the regulation of AI, China may have been driven by an alternate, self-serving, agenda.

D. AI Regulation in the United States

A few steps behind the EU and China, the United States ("US") has been more hesitant to begin regulating AI. The US shying away from making broad, general pronouncements on AI regulation is likely a safeguard against inhibiting the advancement of AI. On May 10, 2018, the White House Summit on Artificial Intelligence for American Industry took place. During this summit, Michael Kratsios, Deputy Assistant to the President for Technology Policy, stated as follows:

> Artificial Intelligence holds tremendous potential as a tool to empower the American worker, drive growth in American industry, and improve the lives of the American people. Our free market approach to scientific discovery harnesses the combined strengths of government, industry, and academia, and uniquely positions us to leverage this technology for the betterment of our great nation.[194]

---

192 *Id.*

193 Zhihao, *supra* note 172.

194 *Summary of the 2018 White House Summit on Artificial Intelligence for American Industry*, WHITE HOUSE OFF. OF SCI. AND TECH. POL'Y (May 10, 2018), https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf.

Kratsio's approach differs slightly from those that have been presented in the EU and China. A free market economy "promotes the production and sale of goods and services, with little to no control or involvement from any central government agency",[195] which is precisely what the EU and China have been trying to prevent. Rather than placing importance on keeping AI safe for consumer use or regulating the standards of AI development, the US is focusing on removing all obstacles[196] that may be in the way of AI advancement. Further, rather than adapting AI to society, the US is taking steps to adapt society to AI. President Donald Trump took executive action[197] to help give American workers the skills to succeed in an economy where AI is ever-growing. To ensure continued federal efforts related to AI, the White House chartered a Select Committee on Artificial Intelligence under the National Science and Technology Council.[198] The goals of the committee are to advise on AI research and development priorities; consider the creation of federal partnerships with industry and academia; improve government coordination of AI research and development; and identify opportunities to use federal data in AI research and development.[199] However, despite this hands off approach he originally presented, Michael Kratsios reflects the US's understanding that US Administration cannot be completely passive. He states that "to realize the full potential of AI for the American people, it will

---

[195] *See* Prateek Agarwal, *Free Market*, INTELLIGENT ECONOMIST (Dec. 9, 2019), https://www.intelligenteconomist.com/free-market/.

[196] WHITE HOUSE OFF. OF SCI. AND TECH. POL'Y, *supra* note 196. *E.g.*, the Department of Transportation released an update to the 2016 Federal Automated Vehicles Policy, providing non-regulatory guidance to automated vehicle developers to enable the safe integration of driverless cars onto American roadways; President Trump signed a Presidential Memorandum to permit States and localities to conduct innovative commercial and public drone operations currently prohibited under FFA regulations; and the FDA approved the first ever AI-based device for medical diagnostics to detect diabetic retinopathy, the leading cause of blindness among working-age Americans.

[197] *Id.* For example, President Trump signed (1) an executive order establishing industry-recognized apprenticeships and creating a cabinet-level task force on apprenticeship expansion; and (2) a presidential memorandum prioritizing high-quality Science, Technology, Engineering, and Math ("STEM") education, with a particular focus on computer science education.

[198] *Id.*

[199] *Id.*

require the combined efforts of industry, academia, and the government."[200] In terms of regulation specifically targeted towards AI, the US has only taken steps into a few subsets of AI thus far. AI regulation in the US has mainly been pursued in the areas of autonomous vehicles[201] and national defense.[202] While both of these areas are important to focus on, as they are large topics within the current discussion of AI, autonomous vehicles and autonomous weapons are only a small part of AI as a whole. AI as a whole is far broader and encompasses more advanced technologies. Thus, it is not a bad starting point to begin regulating in these areas, but the US needs to be aware that the task of regulating AI does not begin and end with these two subcategories.

Since then, the US has been putting forth more of an effort to federally regulate AI. On February 11, 2019, US President Donald Trump issued an Executive Order launching the American AI Initiative.[203] This Executive Order explained that the Federal Government must facilitate AI research and development, as well as create trust surrounding that research and development, prepare citizens for a changing workforce, and protect the nation as a whole.[204] In addition to emphasizing American leadership in AI, this Order stresses the need for cooperation with foreign countries.[205] The American AI Initiative is guided by five principles: (1) driving technological breakthroughs; (2) driving the development of technical standards; (3) training workers to develop and apply AI technologies; (4) protecting American values and creating public trust in AI

---

200 *Id.*

201 *Regulation of Artificial Intelligence: The Americas and the Caribbean*, LIBRARY OF CONGRESS, https://www.loc.gov/law/help/artificial-intelligence/americas.php #us (last visited Jan. 12, 2020). The Department of Transportation is investigating what elements must be considered in drafting regulations for the use of such vehicles, including multi-vehicle convoys, and several states have adopted legislation and regulations for allowing for the testing of autonomous vehicles.

202 *Id.* The Department of Defense has been given the responsibility of crafting policies for the development and deployment of AI systems as they concern national defense.

203 *AI Policy—United States*, FUTURE OF LIFE INST., https://futureoflife.org /ai-policy-united-states/?cn-reloaded=1 (last visited Jan. 12, 2020).

204 *Id.*

205 *Id.*

technologies; and (5) protecting US technological advantage in AI, while also collaborating internationally.[206] Further, all executive departments and agencies involved[207] with AI must follow six strategic objectives.[208] When combined, the five principle initiative and the six part administrative standard both reflect a change in the US's perspective. The US has moved away from wanting to remove all government intervention from the "free market" of AI advancement to providing federal guidelines for AI. Thus, showing an understanding that standards are essential for ensuring that AI can continue to advance unobstructed.

As part of the American AI Initiative, federal agencies have begun to establish guidance for AI development and use. A few examples are (1) the Defense Advanced Research Projects Agency's ("DARPA") Explainable AI program, which "aims to create machine learning techniques that produce more explainable solutions while maintaining high performance and appropriate levels of trust in the system";[209] (2) the National Science Foundation's ("NSF") Program on Fairness in Artificial Intelligence in Collaboration with Amazon, which focuses on funding research on fairness in AI[210]; and (3) DARPA's AI Next Campaign, which will create solutions for defending against potential attacks[211] on AI technologies so that citizens are more likely to place their trust in AI systems. While all of these are worthwhile goals and give off the impression that the US is far along in the process of regulating AI, they simply highlight theoretical milestones that these agencies hope to reach.

---

206 *Id.*

207 *Id.* (*e.g.*, developing or deploying AI, providing educational grants, or regulating or guiding AI).

208 *Id.* ((1) promoting investment of research and development; (2) enhancing access to federal data and resources; (3) reducing barriers to the use of AI; (4) minimizing vulnerability to attacks; (5) training American AI researchers; and (6) implementing a plan to protect US national security interests).

209 *Artificial Intelligence for the American People*, WHITE HOUSE, https://www.whitehouse.gov/ai/ai-american-values/ (last visited Jan. 12, 2020).

210 *Id.*

211 *Id.* (*e.g.*, those that attempt to contaminate training data, modify algorithms, create adversarial inputs, or exploit flaws in AI system goals).

IV. CONCLUSION

Developing a system of regulation for AI is not a simple task, but it is necessary if we want to retain control over developing AI technologies. AI is far more complicated than any technology we have experienced in the past. After all, the main goal of AI is to be able to mimic human-like functioning, but in a way that is superior to human ability. This technology has a mind that is so similar to a human's yet not similar enough that it can be regulated in the same way as we regulate humans. Further, we have never had a technology as diverse and intricate as AI. There are many different subsets and aspects of AI, each tremendously different than the next. Thus, we need to devise a plan for governing AI as a whole, while still putting regulations in place for each specific part—again, something that policymakers have not had to do before. As illustrated by the EU, China, and the US, many countries have taken steps to begin regulating and providing standards for AI. However, many of the proposed solutions are proposed solely in theory, promising only that an agency or committee will create certain rules in the near future. This is certainly still a step in the right direction, but it is one that cannot be guaranteed to foster results until we actually see these rules and regulations being put into effect. Not to mention that once individual countries have created a system of regulation, we must then still figure out a way to make them work together in a global realm.

While not the usual approach, AI is not a usual technology, so maybe the solution is not for each country to develop their own system of regulation independent of each other as has always been done in the past. Due to the complexities of AI, working on a global solution should be made a first priority. Just from examining the EU, China, and the US, it is clear that individual countries all currently have different pieces of the puzzle that they are working towards. The EU is pioneering consumer-friendly, safe, trustworthy, and ethical AI; China is not willing to relinquish the goal of becoming the world leader in AI and has created strategic regulations that provide just enough comfort to the consumer without stripping away developers rights to advance the technology; and the US, taking a slightly more hands off approach, provides a very neutral set of

regulations, weary to lean too far to either side. Every country only has the ability to regulate through its own perspective. Countries will regulate in a way that that will be the most beneficial in the highest number of ways for them specifically. While usually a strategic approach, AI is such a global, interconnected technology that it will never fit within the confines of a one dimensional system of regulation. We need to begin looking at AI through the different lens of each country in order to produce the most carefully tailored solution. AI is an intensely in-depth and complicated technology, and while countries have been making small, productive steps on their own, perhaps a collaborative global effort is what will ultimately allow us to take the leap necessary to create a comprehensive AI regulation plan.