Artificial Intelligence and Public Policy

Adam Thierer, Andrea Castillo O'Sullivan, and Raymond Russell

MERCATUS RESEARCH

MERCATUS CENTER
George Mason University
3434 Washington Blvd., 4th Floor, Arlington, Virginia 22201
www.mercatus.org

Adam Thierer, Andrea Castillo O'Sullivan, and Raymond Russell. "Artificial Intelligence and Public Policy." Mercatus Research, Mercatus Center at George Mason University, Arlington, VA, 2017.

ABSTRACT

There is growing interest in the market potential of artificial intelligence (AI) technologies and applications as well as in the potential risks that these technologies might pose. As a result, questions are being raised about the legal and regulatory governance of AI, machine learning, "autonomous" systems, and related robotic and data technologies. Fearing concerns about labor market effects, social inequality, and even physical harm, some have called for precautionary regulations that could have the effect of limiting AI development and deployment. In this paper, we recommend a different policy framework for AI technologies. At this nascent stage of AI technology development, we think a better case can be made for prudence, patience, and a continuing embrace of "permissionless innovation" as it pertains to modern digital technologies. Unless a compelling case can be made that a new invention will bring serious harm to society, innovation should be allowed to continue unabated, and problems, if they develop at all, can be addressed later.

JEL codes: O38, K20, K24

Keywords: artificial intelligence, machine learning, big data, regulation, permissionless innovation, innovation, precautionary principle

This paper can be accessed at https://www.mercatus.org/publications/artificial-intelligence-public-policy

There is growing interest in the market potential of artificial intelligence (AI) technologies and applications as well as in the potential risks that these technologies might pose. As a result, questions are being raised about the legal and regulatory governance of AI, machine learning, "autonomous" systems, and related robotic and data technologies.

For example, in May 2016, the White House Office of Science and Technology Policy (OSTP) announced four public workshops and requested public comments on "how best to harness the opportunities brought by artificial intelligence."[1] The White House also teed up some of the policy concerns surrounding AI. "Like any transformative technology," the White House noted, "artificial intelligence carries some risk and presents complex policy challenges along several dimensions, from jobs and the economy to safety and regulatory questions."[2]

The method that policymakers ultimately choose to govern the wide range of AI technologies and applications will have a dramatic effect on the ultimate array of opportunities and benefits that could result. Policymakers and regulators face two competing approaches. They can choose to preemptively limit or even ban certain applications out of fears for worst-case scenarios, an option known as the "precautionary principle," or they can prioritize experimentation and collaboration as the default while addressing any issues that do arise as they go, as we call "permissionless innovation."

Many of the comments submitted to the OSTP in that proceeding called for policy interventions of a precautionary nature. Surprisingly, the typical anxieties that have traditionally followed advances in automation technologies—namely, adverse labor market effects—were only a secondary concern of many of the most critical public comments. Instead, the specters of discrimination and structural social inequality catalyzed the most prohibitory policy recommendations.

1. Ed Felten, "Preparing for the Future of Artificial Intelligence," *White House* blog, May 3, 2016.
2. Ibid.

> “We will discuss the profound economic benefits that AI technologies may provide to the United States and then examine how overly prohibitive regulatory proposals could inadvertently undermine this promising industry before it has a chance to develop.”

For example, several respondents warned that the threat of “algorithmic bias” against protected classes could necessitate oversight by public-private ethics bodies or redress by other government mechanisms to prevent and punish disparate impact.[3]

Indeed, the White House released a report on big data and machine learning during the same month that the OSTP sent out its request for comment. Although the report did praise these technologies for the opportunities they could create to overcome social inefficiencies, it unfortunately advocates a mandate on programmers called “equal opportunity by design.”[4] This principle would require technologists to design “data systems that promote fairness and safeguard against discrimination from the first step of the engineering process and continuing throughout their lifespan”—thereby potentially granting regulators enormous control over each step of the testing and development process for computer code. A requirement that innovators receive constant permission from a number of different regulators before deploying new algorithms could ultimately backfire against the report’s noble goals by slowing innovation and preventing the many promising technologies that its authors believe could alleviate inequality.[5]

In this paper, we recommend a different policy framework for AI technologies. At this nascent stage of

---

3. “Preparing for the Future of Artificial Intelligence: In Response to Office of Science and Technology Policy,” Center for Democracy and Technology, July 27, 2016; Shannon Vallor, “On Artificial Intelligence and the Public Good,” Markkula Center for Applied Ethics at Santa Clara University, July 19, 2016.
4. Cecilia Muñoz, Megan Smith, and DJ Patil, *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*, Executive Office of the President, May 2016.
5. An overly strict liability regime can likewise dull innovation. The report also favorably cites a 2014 White House report calling on the Justice Department, Consumer Financial Protection Bureau, and Equal Employment Opportunity Commission to “develop a plan for investigating and resolving” any “big data analytics that have a discriminatory impact on protected classes.” John Podesta et al., *Big Data: Seizing Opportunities, Preserving Values*, Executive Office of the President, May 2014.

AI technology development, we believe a better case can be made for prudence, patience, and a continuing embrace of permissionless innovation as it pertains to modern digital technologies. Permissionless innovation refers to the idea that "experimentation with new technologies and business models should generally be permitted by default. Unless a compelling case can be made that a new invention will bring serious harm to society, innovation should be allowed to continue unabated, and problems, if they develop at all, can be addressed later."[6]

Policymakers initially may be tempted to preemptively restrict AI technologies out of an abundance of caution for the perceived safety, welfare, and market risks these new innovations might seem to pose. However, an examination of the history of early US internet policy suggests that these concerns can be adequately addressed without needlessly quashing a potentially revolutionary new industry before it even has the chance to develop.

Accordingly, this paper will compare the advent of AI technologies to the development of the commercial internet in the 1990s to provide insight into how policymakers may champion pro-growth policies while maintaining an appropriate level of oversight and accountability for consumers. We will discuss the profound economic benefits that AI technologies may provide to the United States and then examine how overly prohibitive regulatory proposals could inadvertently undermine this promising industry before it has a chance to develop.

We will conclude by outlining an alternative regulatory path for AI technologies founded on the principles of patience and permissionless innovation. In particular, policymakers should prioritize developing an appropriate understanding of the varied sector of artificial intelligence technologies from the outset and developing an appreciation for limitations of our ability to forecast either future AI technological trends or crises that may ultimately fail to materialize.

This policy approach—rooted in humility, flexibility, and forbearance—will help ensure that policies will promote both innovation and the public good.

POTENTIAL BENEFITS OF AI

The phrase "artificial intelligence" is an umbrella term for technologies that appear to act as if they were rational beings.[7] Employing some combination of

---

6. Adam Thierer, *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom*, 2nd ed. (Arlington, VA: Mercatus Center at George Mason University, 2016).
7. A. M. Turing, "Computing Machinery and Intelligence," *Mind* 59 (1950): 433, 442.

algorithms, "machine learning,"[8] sensory feedback systems, and automation, AI technologies simply perform the functions that they were programmed to learn. Yet, to the human observer, it appears as if a machine or computer did not perform the functions at all. It seems as if truly a human is behind the scenes. A human is not, but this clever kind of programmed labor can provide considerable benefits to society.

Applications

Examples of promising AI technologies include the following:[9]

- Artificial neural networks that assist humans in diagnosing medical ailments and recommending treatments,[10] identifying fraud or error in markets,[11] guiding manufacturing processes,[12] and even producing consumer goods.[13]
- Visual-spatial recognition systems that allow computers to analyze and interpret images and videos. These technologies can be employed to train medical students,[14] develop car-safety systems,[15] and generate street-view maps.[16]

---

8. Machine learning is a method of programming that allows computers to build their own analytical models. For more information, see Alex Smola and S. V. N. Vishwanathan, *An Introduction to Machine Learning* (Cambridge, UK: Cambridge University Press, 2008).
9. These technologies are discussed further in Nicholas Chen et al., "Global Economic Impacts Associated with Artificial Intelligence" (Analysis Group study for Facebook, 2016), 23.
10. Indeed, the White House launched its own "Precision Medicine Initiative" to promote just these applications in 2015. For more background on medical neural network technologies, see Filippo Amato et al., "Artificial Neural Networks in Medical Diagnosis," *Journal of Applied Biomedicine* 11, no. 2 (2013).
11. Adam Fadlalla and Chien-Hua Lin, "An Analysis of the Applications of Neural Networks in Finance," *Interfaces* 31, no. 4 (2001).
12. Martin Röscheisen, Reimar Hofmann, and Volker Tresp, "Neural Control for Rolling Mills: Incorporating Domain Theories to Overcome Data Deficiency," in *Advances in Neural Information Processing Systems*, vol. 4, ed. J. E. Moody, S. J. Hanson, and R. P. Lippmann (San Mateo, CA: Morgan Kaufmann, 1992).
13. Kazuo Asakawa and Hideyuki Takagi, "Neural Networks in Japan," *Communications of the ACM* 37, no. 3 (1994).
14. Mary Hegarty et al., "The Role of Spatial Cognition in Medicine: Applications for Selecting and Training Professionals," in *Applied Spatial Cognition: From Research to Cognitive Technology*, ed. Gary L. Allen (Mahwah, NJ: Lawrence Erlbaum Associates, 2007).
15. Naomi Tajitsu, "Toyota to Build Artificial Intelligence-Based Driving Systems in Five Years," *Reuters*, June 20, 2016.
16. Dave Gershgorn, "Facebook Made Detailed Maps of 20 Countries for Its Internet Drones," *Popular Science*, February 22, 2016.

- “Virtual private assistants” that learn from user behavior to help recommend activities and keep track of obligations.[17]
- Automation technologies that enable machines to operate without human assistance by learning from available data. Examples include automated vehicles,[18] instruments for scientific research,[19] and tools of data analysis.[20]

This list makes it clear that a wide range of promising and distinct technologies and services fall under the umbrella of artificial intelligence. One of the first tasks facing policymakers examining the broad area of AI technologies is to outline a clear and appropriate system of definitions for these technologies with the input of industry and academic researchers.

This is no small task. Indeed, some of the most seasoned artificial intelligence experts struggle to formulate a concise definition and taxonomy of these technologies. The difficulty is due partially to the ephemeral nature of the technology itself and partially to the uneven history of human interest and understanding in this subject.

The term *artificial intelligence* has assumed many connotations during its history, enduring periods of skepticism and pessimism and now enjoying a renewal of interest. Progress in artificial intelligence research has been punctuated by several so-called AI winters, periods during which public interest, investment, and demand collapsed. Twice—once in the 1970s and again a decade later—public- and private-sector investment in AI slowed as researchers failed to deliver sufficiently consequential results. In the years leading up to these “winters,” governments and investors were exuberant about the possibilities of intelligent systems, and many tended to overstate the progress of AI research, thus leading to public discouragement and stagnation when reality caught up.

A certain stigma became associated with the term artificial intelligence, so researchers in AI subfields tended to refer to their projects by more specific names. Neural networks, machine learning, and data mining are terms that all belong in the basket of AI. Systems that performed well at only very specific tasks were dismissed and not considered AI,[21] but as late as 2007 the *Economist* lamented that many investors still associated the term *artificial intelligence*

17. Richard Waters, “Artificial Intelligence: A Virtual Assistant for Life,” *Financial Times*, February 22, 2015.
18. Erico Guizzo, “How Google’s Self-Driving Car Works,” *IEEE Spectrum*, October 18, 2011.
19. Lizzie Buchen, “Robot Makes Scientific Discovery All by Itself,” *Wired*, April 2, 2009.
20. Patrick Laurent, Thibault Chollet, and Elsa Herzberg, “Intelligent Automation Entering the Business World,” *Inside* (Deloitte) 8 (2015).
21. James Hendler, “Is Artificial Intelligence Headed for Another AI Winter? Not If We Take Action Now,” *IEEE Intelligent Systems* 23, no. 2 (2008).

with failure and underperformance.[22] Because of the dynamic history of artificial intelligence, the field comprises a web of mutable terms—boundaries between disciplines are perpetually evolving and often difficult to discern. The following is a brief glossary of useful terms:

- *Algorithm*. A series of discrete, conditional instructions. In computing, algorithms enumerate a list of operations to carry out. Much as a recipe book gives directions to a chef, an algorithm informs a computer of the steps it must take to deliver a desired result.
- *Artificial intelligence*. The exhibition of intelligence by a machine. An AI system is capable of undertaking high-level operations; AI can perform near, at, or beyond the abilities of a human. This concept is further divided into weak and strong AI.
- *Autonomous robotic system*. A tangible application of AI technology. Most AI technology in use today is intangible—computer programs that can execute financial transactions and filter spam out of an inbox. Increasingly, the world is seeing more tangible applications in everyday life. Robotic vacuum cleaners are an example of fully autonomous robotic systems. The SpaceX oceangoing landing platform,[23] surgery robots,[24] and unmanned aerial military vehicles[25] are examples of systems that have some autonomous capabilities today and may become fully autonomous in the coming decades.
- *Big data*. Any dataset consisting of a sufficiently large number of observations that a human cannot analyze it unaided by computers. Data scientists can employ machine learning models to establish patterns and provide predictions. A human, for example, can read a few books every month and augment his or her understanding of the world. But the reader cannot hope to read every book ever published; a machine can, in a relatively short amount of time, process more information than a human could absorb in a lifetime.
- *Deep learning*. A class of machine learning techniques involving several "layers" of abstraction. The popular press often uses the term to refer to more general machine learning techniques, but in reality deep learning

22. "Are You Talking to Me?," *Economist*, June 7, 2007.
23. "X Marks the Spot: Falcon 9 Attempts Ocean Platform Landing," SpaceX website, December 16, 2014.
24. Eliza Strickland, "Autonomous Robot Surgeon Bests Humans in World First," *IEEE Spectrum*, May 4, 2016.
25. Lora G. Weiss, "Autonomous Robots in the Fog of War," *IEEE Spectrum*, July 27, 2011.

is a set of specific methods and algorithms. Many neural networks are deep learning systems; there are multiple steps taken between input and output during which "neurons" interact. Digital image classification is a common application of deep learning methods. For example, the computer is fed an image that is meaningful to humans, the "visible layer." The computer then progressively identifies features that become more abstract in the "hidden layers," before an output is reached that is again meaningful to humans.[26]

- *Machine learning*. The process by which a computer can train and improve an algorithm or model without step-by-step human involvement. One of the simplest classes of machine learning algorithms is linear regression. In this case, a computer interpolates a linear relationship between known input and output data. The error-minimizing outcome is the machine's best attempt at a prediction of future observations. An associated concept, data mining, refers to the process by which machines can "learn," generating useful conclusions and models by analyzing vast datasets.
- *Neural networks*. A class of machine learning systems that draws inspiration from neural functions in biology. A neural network is composed of many neurons—nodes for information processing that receive inputs from and send outputs to other nodes.
- *Weak AI*. Today, AI is generally "weak," optimized for specific, narrowly defined tasks. These systems include automated financial trading programs and Siri, the iPhone's personal assistant. Although both perform very well in the limited set of applications for which they were designed, weak AI systems are not genuinely "intelligent" in the human sense of the

26. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (Cambridge, MA: MIT Press, 2016), 6–8.

"Weak AI systems are not genuinely 'intelligent' in the human sense of the word. In contrast, a 'stronger' AI system might perform competently in several different fields."

word. In contrast, a "stronger" AI system might perform competently in several different fields.[27]

Most artificial intelligence implemented in the next few decades will likely remain weak and highly specialized; these systems are not useful for applications unrelated to their intended purpose. For example, a piece of computerized trading software cannot navigate the best route through rush-hour traffic or suggest three Chinese restaurants nearby. Siri cannot make split-second decisions between buying ounces of gold or shares of Goldman Sachs. To understand the difference between strong and weak AI, take the latter case of a personal assistant. Siri, Cortana, and Alexa fall into the weak category, while a strong version might outperform a human assistant in many fields simultaneously.

The fundamental driver of current AI systems is data. To illustrate the integrality of good data in successful AI systems, we will examine a canonical problem that is routinely used in university-level machine learning courses. For a long time, algorithms and the machines on which they ran were insufficiently robust to identify basic objects. To a human, image recognition is trivial. Very early in life, a child is already able to discern the number eight from the family cat. The basic principle of biological knowledge acquisition is similar to that of a computer; a child learns what is an eight and what is a cat by seeing both repeatedly and, over time, assigning definitions to these concepts. Human learning is much like machine learning in that it is data driven.

Unlike humans, however, machines lack the intuition to make such inferences. The first artificial neural network capable of reliably recognizing a zip code came about in the early 1980s, and it took several more decades before a computer could determine that a given photo contained a cat. The process that comprises such a classifier is as follows: data (e.g., an image of a cat) are input in a machine-readable format, the machine uses an algorithm to test the image against previous experience, and a conclusion that can be read by a human is output.

A computer does not instinctively know what an animal is or what distinguishes an animal from a number. This distinction can be taught to a machine through training that involves large amounts of data. Gradually, a computer program learns to distinguish between numbers and animals; with much more data, it is even possible for a computer to separate housecats from big cats. To a human, this task is not difficult, as most people can tell the difference between a tabby and a tiger. But a machine does not enjoy this privilege, and when presented with

27. Conner Forrest, "Mini-Glossary: AI Terms You Should Know," *Tech Republic*, August 5, 2015.

photos of both, it sees at first only two brown masses that resemble each other more than do an animal and a number.

A programmer may be able to build a reliable artificial intelligence system that can, using neural networks and a large amount of training data, identify the subtle differences between breeds of housecat. The resulting classifier is a stronger AI than the zip code–recognizing program. A hierarchy of sophistication emerges as we include more systems; the aforementioned applications are all examples of weak or "narrow" AI—navigation systems, spam filters, and game-playing machines like Alpha Go fall into this category. More highly developed artificial general intelligence (AGI), if developed, would have features similar to the reasoning faculties of a human, although the system's architecture might not resemble that of the human brain.[28] The principal characteristics that would define AGI systems include the ability to become proficient in a variety of fields and to do so at least somewhat autonomously.[29] In the more distant future, some speculate that an even more advanced system, artificial superintelligence (ASI), will emerge; ASI will exceed the collective intelligence of humanity.[30] The development of artificial superintelligence technology is not guaranteed. Sophisticated design and implementation challenges would have to be resolved. Even artificial general intelligence is by no means inevitable; for many decades, researchers have been grappling with the problems of fabricating cognition.[31]

---

28. See Ben Goertzel and Cassio Pennachin, "The Novamente Artificial Intelligence Engine," in *Artificial General Intelligence*, ed. Ben Goertzel and Cassio Pennachin (New York: Springer, 2007). The authors present a table of characteristics that differentiate narrow AI from more capable, generally intelligent systems. Relative to AGI systems, a narrow AI may require a programmer to possess substantial domain knowledge about the task at hand, is less able to improve on its own, and may not produce outputs easily understood by laypeople.
29. Peter Voss, "Essentials of General Intelligence: The Direct Path to Artificial General Intelligence," in Goertzel and Pennachin, ed., *Artificial General Intelligence*.
30. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014); Jerry Kaplan, *Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence* (New Haven, CT: Yale University Press, 2015).
31. For some of the earliest works on this subject, see John McCarthy, "Programs with Common Sense," *Mechanization of Thought Processes*, vol. 1 (London: Her Majesty's Stationery Office, 1959); John McCarthy and Patrick J. Hayes, "Some Philosophical Problems from the Standpoint of Artificial Intelligence [1969]," repr. in *Readings in Artificial Intelligence*, ed. B. L. Webber and N. J. Nilsson (Los Altos, CA: Morgan Kaufmann Publishers Inc., 1981). For an example of an approach to building such a system, see Don Perlis et al., "A Broad Vision for Intelligent Behavior: Perpetual Real-World Cognitive Agents," in *2013 Annual Conference on Advances in Cognitive Systems: Workshop on Metacognition in Situated Agents*, ed. D. Josyula, P. Robertson, and M. T. Cox (College Park: University of Maryland, 2013).

Currently, interactions between humans and artificially intelligent systems are commonplace. Probably the most frequent use involves querying a search engine; when a searcher inputs text into a search service, an algorithm weights websites by significance and relevance, generating a list of results that the user will likely find interesting.[32] As recently as January 2016, 15 percent of Google searches were terms completely new to the engine.[33] Search engines have come to rely on deep learning–based processes. Google, Facebook, and Microsoft have rolled out deep learning–based search algorithms for general use, whereas WolframAlpha has developed a powerful, highly specialized AI platform that performs muscularly against computational and science-related queries.[34]

In other settings, AI is already improving the lives of many people. In August 2015, IBM acquired Merge Technology, a firm that has amassed a database of billions of medical images. Using Watson, the IBM AI project that famously appeared on *Jeopardy* in 2011, the company intends to improve computer recognition of various medical conditions. IBM will use its vast new dataset to train Watson using machine learning;[35] the Watson algorithm will, the company hopes, be robust enough to rely on in diverse applications.[36] The predictive power of Watson is not intended to replace radiologists, oncologists, and internists, but these efforts should augment their diagnostic and treatment capabilities.

Thanks to advancements in AI, a person who is deaf can more easily interact with the world. The internet has dramatically increased human interconnectivity. Services such as Skype, FaceTime, and internet-based audio telephony have improved long-distance communication, making it possible for two people on opposite sides of the world to interact nearly as closely as they would in the same room. But communication platforms alone do not yield benefits to the deaf and hard-of-hearing community. As AI systems have improved and technological adoption has increased, the internet has become more inclusive of those unable

32. Google PageRank, for example, measures the importance of a page by the number and significance of other websites that link to it. See Tahseen A. Jilani et al., “A Survey and Comparative Study of Different PageRank Algorithms,” *International Journal of Computer Applications* 120, no. 24 (2015).
33. Greg Notess, “Search Engine Update,” *Online Searcher* 40, no. 3 (2016): 8.
34. Nathan Sikes, “Deep Learning and the Future of Search Engine Optimization,” *TechCrunch*, June 18, 2015.
35. A technical example of a diagnostic medical algorithm is the Markov process concept. The product of this approach has outperformed human-only healthcare provision in both cost and successful treatment rates. See, for example, Casey Bennett and Kris Hauser, “Artificial Intelligence Framework for Simulating Clinical Decision-Making: A Markov Decision Process Approach,” *Artificial Intelligence in Medicine* 57, no. 1 (2013).
36. Robert McMillan and Elizabeth Dwoskin, “IBM Crafts a Role for Artificial Intelligence in Medicine,” *Wall Street Journal*, August 11, 2015.

to communicate through voice alone. A combination of speech and sign recognition has extended the possibilities of video calling. This application is at the intersection of augmented reality, natural language processing, and computer vision machine learning. For several decades, researchers have worked to build systems that can recognize, interpret, and transcribe sign language.[37]

Recently, advanced machine learning concepts have been applied to sign language gesture recognition, improving communication between signers and nonsigners. Artificial neural networks can recognize individual features that compose the overall hand and construct a complete picture of a given sign, inferring its meaning. Genetic algorithms accomplish the same goal by beginning with an approximate test case that resembles the sign captured on the computer's webcam. The algorithm improves and fine-tunes itself until it reaches a desired result.[38] In Spain, researchers have developed a framework for greater accessibility on public transit systems. A person who is deaf interacting with a bus ticket seller could ask a question using sign language, the question could be translated into a spoken language and synthesized, and the bus system employee would hear a computer-generated interpretation. Similarly, the employee's answer would be rendered into sign language by an avatar on a screen outside the kiosk.[39]

Researchers interested in accessibility could also employ machine-learning strategies to benefit those with learning disabilities. Although there are more than 5 million pages on the English Wikipedia, many people may find the online encyclopedia's level of technicality prohibitive. The global community of Wikipedia contributors has made steady progress toward compiling a Simple English version using a dictionary of only 2,000 words. Free of jargon, this resource is more accessible and useful to those who are new to the English language or have reading disabilities. The Simple English Wikipedia currently comprises just over 120,000 articles—only a small fraction of the content available in standard English. Until recently, the only way to generate coherent, simplified articles was by hand.[40]

Those efforts were not in vain; manual simplification is an important step that will enable robust AI systems. Construction of a successful text simplifier

---

37. Becky Sue Parton, "Sign Language Recognition and Translation: A Multidisciplined Approach from the Field of Artificial Intelligence," *Journal of Deaf Studies and Deaf Education* 11, no. 1 (2006).
38. Ankit Chaudhary et al., "Intelligent Approaches to Interact with Machines Using Hand Gesture Recognition in Natural Way: A Survey," *International Journal of Computer Science and Engineering Survey* 2, no. 1 (2011).
39. Veronica López-Ludeña et al., "Translating Bus Information into Sign Language for Deaf People," *Engineering Applications of Artificial Intelligence* 32 (2014).
40. For a comprehensive overview of the potential sector-by-sector benefits of AI, see Peter Stone et al., "Artificial Intelligence and Life in 2030" (One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel, Stanford University, Stanford, CA, September 6, 2016), 18–41.

cannot exist in a vacuum because a large amount of data is required. Researchers have thought to employ the corpus of the Simple English Wikipedia to train AI text simplifiers. Some current efforts make use of neural networks[41] and decision trees[42] to create fluent sentences. Of course, obtaining strong training data is an important concern for anyone building an AI system, so researchers are looking to existing simplification databases and crowdsourcing[43] for assistance in building and testing their models.[44]

Meteorology also benefits from the predictive power of data-intensive AI. Current and near-term AI applications in weather forecasting will improve safety and efficiency in transportation. Using patterns elucidated from large amounts of existing data, weather-predicting neural networks can generate robust predictions about the future.[45] As long as two decades ago, researchers had great success in developing precipitation forecasts using neural networks.[46] This modeling finds uses in effective flood mitigation, hydroelectric power station safety, and crop yield forecasting. In a similar vein, neural networks are useful for solar weather forecasting. Using machine learning, a team at the University of Zagreb in Croatia created a strong model for travel times of coronal mass ejections;[47] these disturbances generate streams of charged particles that, when interacting with the Earth's magnetic field, cause geomagnetic storms. Coronal mass ejections pose particular risks to modern life because many electrical systems on Earth are susceptible to the electromagnetic disturbances they cause.

---

41. Tong Wang et al., "Text Simplification Using Neural Machine Translation," in *AAAI '16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (Palo Alto, CA: Association for the Advancement of Artificial Intelligence, 2016).
42. Gustavo H. Paetzold and Lucia Specia, "Text Simplification as Tree Transduction," in *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology* (Fortaleza, CE, Brazil: Sociedade Brasileira de Computação, 2013).
43. Walter S. Lasecki, Luz Rello, and Jeffrey P. Bigham, "Measuring Text Simplification with the Crowd," in *Proceedings of the 12th Web for All Conference* (New York: Association for Computing Machinery, 2015).
44. Sanja Štajner, Ruslan Mitkov, and Horacio Saggion, "One Step Closer to Automatic Evaluation of Text Simplification Systems," in *Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations* (Stroudsburg, PA: Association for Computational Linguistics, 2014).
45. Andrew Culclasure, "Using Neural Networks to Provide Local Weather Forecasts" (master's thesis, Georgia Southern University, 2013).
46. Tony Hall, Harold E. Brooks, and Charles A. Doswell III, "Precipitation Forecasting Using a Neural Network," *Weather and Forecasting* 14 (June 1999).
47. Davor Sudar, Bojan Vršnak, and Mateja Dumbović, "Predicting Coronal Mass Ejections Transit Times to Earth with Neural Network," *Monthly Notices of the Royal Astronomical Society* 456, no. 2 (2016).

This survey only scratches the surface of the wide range of potential AI applications and benefits. For example, a recent report from the Information Technology & Innovation Foundation's Center for Data Innovation documented 70 examples of how "AI is driving innovation in the public and private sectors, generating substantial social and economic value, and transforming everyday life around the globe."[48]

Market Potential of AI

The American technology sector, long the envy of the world for spearheading revolutionary new technologies for consumers and business, has invested considerable capital in the future of AI. Private investment in AI technologies has grown substantially in recent years, from $1.7 billion in 2010 to $14.9 billion in 2014.[49] Venture Scanner reports an explosion of startup activity related to AI technologies in recent years, growing from fewer than 20 AI startups founded in 2003 to 188 startups founded in 2013 alone.[50] In 2014 and 2015, tech giants such as Google, Microsoft, Apple, Amazon, IBM, Yahoo!, Facebook, and Twitter made at least 26 major acquisitions of AI and machine learning firms, totaling roughly $5 billion.[51] A 2013 study by analysts at Booz & Company found that the healthcare industry invested the most in artificial intelligence technologies, with $13.8 billion in investment that year, followed by electronics manufacturers with $9.7 billion and software and web companies with $7.7 billion in investment.[52]

Accordingly, the economic benefits of AI are projected to be substantial. One recent study used benchmarks derived from methodologically conservative studies of broadband internet, mobile phones, and industrial robotics to estimate that the economic impact of AI could be between $1.49 trillion and $2.95 trillion over the next 10 years.[53] With less strict assumptions, the economic benefits

---

48. Daniel Castro and Joshua New, "The Promise of Artificial Intelligence: 70 Real-World Examples" (Information Technology and Innovation Foundation, Center for Data Innovation, Washington, DC, 2016).
49. Chen et al., "Global Economic Impacts."
50. "Artificial Intelligence Companies Founded by Year—Q4 2016," Venture Scanner, November 30, 2016.
51. Ibid.
52. Gitta Rohling, "Facts and Forecasts: Boom for Learning Systems," Innovations newsletter (Siemens), October 1, 2014.
53. Ibid., 23. "Growth in AI producing sectors could lead to increased revenues, and employment within these existing firms, as well as the potential creation of entirely new economic activity. Productivity improvements in existing sectors could be realized through faster and more efficient processes and decision making as well as increased knowledge and access to information." Chen et al., "Global Economic Impacts Associated with Artificial Intelligence," 3.

“Autonomous or ‘driverless’ car technologies could save the United States $1.3 trillion in annual costs, or 8 percent of annual GDP, and $5.6 trillion globally once those technologies have fully penetrated.”

could be greater still. Another report from Transparency Market Research anticipates that the global artificial intelligence market, which was valued at $126.24 billion in 2015, will grow to reach a value of $3.1 trillion by 2024.[54]

Other studies seek to quantify the anticipated economic benefits of specific AI applications. A 2013 report by analysts at Morgan Stanley projected that autonomous or “driverless” car technologies could save the United States $1.3 trillion in annual costs, or 8 percent of annual GDP, and $5.6 trillion globally once those technologies have fully penetrated.[55] Productivity gains from people spending less time stuck in traffic in their cars could reach $647 billion each year. Fuel cost savings could be as high as $168 billion per year, and another $488 billion in savings could accrue from traffic accident and fatality avoidance.

AI systems are also expected to yield considerable economic benefits for one of the world’s fastest-growing and critical industries: healthcare and medical research. “Smart diagnosis” and patient tracking tools are expected to empower healthcare providers and patients alike to make better care decisions at lower cost.[56] These tools could improve population health, and therefore economic productivity. Thus, industry analysts anticipate that AI health technologies will be a big market success. A 2015 report from Frost & Sullivan calculated that the cognitive computing technology market would grow from $633.8 million in earned revenues in 2014 to more than $6.6 trillion in 2021.[57]

Advances in robotics, too, are projected to yield considerable cost savings and productivity gains across a broad

54. “Artificial Intelligence Market—Global Industry Analysis, Size, Share, Growth, Trends and Forecast 2016–2024” (Market Transparency Research, Albany, NY, March 2, 2016).
55. Ravi Shanker et al., “Autonomous Cars: Self-Driving the New Auto Industry Paradigm” (Morgan Stanley Blue Paper, Morgan Stanley Research, New York, 2013).
56. Andy Kessler, “Siri, Am I About to Have a Heart Attack?,” *Wall Street Journal*, January 9, 2017.
57. Harpreet Singh Buttar and Venkat Rajan, *Cognitive Computing and Artificial Intelligence Systems in Healthcare* (Santa Clara, CA: Frost & Sullivan, 2015).

swath of industries. McKinsey and Company reports that the global economic effect of advanced robotics could range from $1.7 trillion to $4.5 trillion by 2025, with $800 billion to $2.6 trillion worth of value derived from healthcare applications of robotics such as machine-assisted precision surgery and robotic prosthetics, $700 billion to $1.4 trillion accruing owing to improvements in manufacturing, and $200 billion to $500 billion in savings from household production.[58]

Artificial intelligence technologies have been developed and deployed to augment manufacturing techniques for several decades.[59] AI applications can be employed to design factory equipment,[60] optimize processes,[61] manage inventories, maintain factory equipment, and assure product quality.[62] New advances in AI techniques have spurred an uptick in manufacturing research and development and integration. A survey by KPMG finds that 25 percent of respondents have already invested in AI and robotic technologies for manufacturing.[63] Another 40 percent have major investments in these techniques planned for the next two years. Analysts with Bank of America and Merrill Lynch report that robotics and AI will be performing 45 percent of manufacturing tasks by 2025, compared with 10 percent today.[64] The value of the productivity gains and saved labor costs yielded by industrial robotics and AI techniques could be as high as $1.2 trillion in 2025.

These technologies can also deliver cost savings for municipal service providers and governments.[65] "Smart city" technologies embed connected sensors in physical municipal systems to provide instant updates of movements and use them to better deliver products and services and cut down on waste.[66]

---

58. James Manyika et al., *Disruptive Technologies: Advances That Will Transform Life, Business, and the Global Economy* (San Francisco: McKinsey Global Institute, May 2013).
59. Farid Meziane et al., "Intelligent Systems in Manufacturing: Current Developments and Future Prospects," *Integrated Manufacturing Systems* 11, no. 4 (2000).
60. Nianyi Chen, Conghe Li, and Pei Qin, "KDPAG Expert System Applied to Materials Design and Manufacture," *Engineering Applications of Artificial Intelligence* 11, no. 5 (1998).
61. H.-C. Zhang and S. H. Huang, "Applications of Neural Networks in Manufacturing: A State-of-the-Art Survey," *International Journal of Production Research* 33, no. 3 (1995).
62. Peter Kopacek, "Intelligent Manufacturing: Present State and Future Trends," *Journal of Intelligent and Robotic Systems* 26, no. 3 (1999).
63. Doug Gates, Tom Mayor, and Erich L. Gampenrieder, "Competing for Growth: How to be a Growth Leader in Industrial Manufacturing" (KPMG Global Manufacturing Outlook, KPMG International, Amsterdam, 2016).
64. Beijia Ma, Sarbjit Nahal, and Felix Tran, "Robot Revolution: Global Robot and AI Primer," *Thematic Investing* newsletter, December 16, 2015.
65. Conor Griffin et al., "Advanced Science and the Future of Government" (World Government Summit Thought Leadership Series report, Economist Intelligence Unit, London, 2016).
66. Andries van Dijk et al., "Smart Cities: How Rapid Advances in Technology Are Reshaping Our Economy and Society" (Deloitte, Amsterdam, November 2015).

For example, an AI-assisted energy system could learn about patterns of peak and low use to optimize energy use and save money. One industry analysis projects that the value of the smart cities market will grow from $312 billion in 2015 to $757.7 billion by 2020.[67]

Of course, as the adage goes, predictions are difficult to make, especially about the future. Poor regulatory frameworks or heretofore unseen technical challenges could limit the full potential of these exciting applications. Or, more optimistically, researchers could discover new applications that exceed even the most bullish of the discussed projections. In that case, the optimistic economic projections also generate uncertainties about the future of employment and compensation, as we will soon discuss. The technologies should succeed or fail on their own merits, not because of poorly designed policies. However, as we will now discuss, some critics of artificial intelligence technologies would prefer that the government preemptively clamp down on these technologies for fear of worst-case scenarios.

REGULATORY THREATS TO AI INNOVATION

Many people's understanding of artificial intelligence technologies is unfortunately largely informed by dystopian science fiction movies, television, and books.[68] If one's sole conception of a technology comes from Hollywood depictions of killer robotic systems run amok or Huxleyesque synthetic pleasure prisons, it is understandable that one might want to use the force of regulation to clamp down decisively on this "threat." But these fictional representations are just that: fictional. AI technologies are at the same time much more benign and much more fantastic in real life.

Yet it is not unusual to also hear those same end-of-the-world dystopian scenarios thrown around in many nonfiction books and essays. *Terminator*-inspired tales of killer robots destroying humanity are a common feature of many of the most popular works on AI and robotics.[69]

67. "Smart Cities Market by Solution and Services for Focus Areas (Transportation—Rail & Road, Utilities—Energy, Water, & Gas, Buildings—Commercial & Residential, and Smart Citizen Services—Education, Healthcare, & Security): Global Forecast to 2020" (Report no. TC 3071, Markets and Markets Research, Pune, India, and Seattle, May 2016).
68. Dominic Basulto, "Can We Just Stop with All These Tech Dystopia Stories?," *Washington Post*, December 8, 2015.
69. For examples, see Ben Austen, "The *Terminator* Scenario: Are We Giving Our Military Machines Too Much Power?," *Popular Science*, January 13, 2011; John Markoff and Claire Cain Miller, "As Robotics Advances, Worries of Killer Robots Rise," *New York Times*, June 16, 2014.

The general public, therefore, may be predisposed to fear AI technologies and applications more on the basis of fiction than on fact. But expert commentators have issued doomsday predictions more tethered in reality as well.[70] Such fears have prompted some critics to make the case for a preemptive government regulation of AI technologies. The rationales for control are varied and include concerns ranging from deindustrialization to dehumanization[71] as well as worries about the "fairness" of the algorithms behind AI systems.[72] Some of those specific concerns are discussed later in this paper.

Because of the assorted anxieties associated with AI, some argue that policymakers should "legislate early and often" to "get ahead of" the hypothetical problems.[73] Specifics are often in short supply, with some critics simply hinting that "something must be done" to address amorphous concerns.[74]

Other scholars have provided more concrete regulatory blueprints, however. They propose, among other things, the passage of broad-based legislation[75] such as an Artificial Intelligence Development Act,[76] as well as the creation of a federal AI agency[77] or possibly a Federal Robotics Commission[78] or National Algorithmic Technology Safety Administration.[79] Those proposed laws and agencies would establish a certification process requiring innovators to subject their technologies to regulatory review to "ensure the safety and security of their A.I."[80] Or, at a minimum, such agencies would advise other federal, state, and local officials and organizations on how to craft policy for AI and robotics.

---

70. For example, in Superintelligence, Nick Bostrom outlines potential pathways for AI technologies that would have beneficial or detrimental outcomes depending on their designs and implementations.
71. Nicholas Carr, *The Glass Cage: Automation and Us* (New York: W. W. Norton & Company, 2014); Jerry Kaplan, *Humans Need Not Apply*, 7. (Kaplan suggests that AI systems "can wreak havoc on an unimaginable scale in the blink of an eye.")
72. Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge, MA: Harvard University Press, 2015).
73. John Frank Weaver, "We Need to Pass Legislation on Artificial Intelligence Early and Often," *Slate*, September 12, 2014.
74. Samantha Shorey and Philip N. Howard, "Automation, Big Data, and Politics: A Research Review," *International Journal of Communication* 10 (2016).
75. Alex Rosenblat, Tamara Kneese, and danah boyd, "Understanding Intelligent Systems" (Open Society Foundations' Future of Work commissioned research paper, Data & Society Research Institute, New York, October 8, 2014), 11.
76. Matthew U. Scherer, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies," *Harvard Journal of Law and Technology* 29, no. 2 (2016): 393–97.
77. Ibid, 395–97.
78. Ryan Calo, "The Case for a Federal Robotics Commission" (Brookings Institution, Washington, DC, September 2014).
79. Andrew Tutt, "An FDA for Algorithms," *Administrative Law Review* 69, no. 1 (2017).
80. Scherer, "Regulating Artificial Intelligence Systems," 394.

Proposals that grant regulators broad prohibitory authorities are built on precautionary principle–based reasoning. The precautionary principle generally refers to the belief that new innovations should be curtailed or disallowed until their developers can prove that they will not cause any harms to individuals, groups, specific entities, cultural norms, or various existing laws, norms, or traditions.

It is certainly true that AI technologies could give rise to some of the problems that critics suggest. And we should continue to look for constructive solutions to the potentially thorny problems that some of these critics suggest are coming. That does not mean that top-down, technocratic regulation is sensible, however.

Automation and the Future of Work

Concerns have been growing about the "rise of the robots" and the impact of automation and AI adoption on the present-day workforce.[81] But these fears are older than many may realize. A *LIFE Magazine* cover from July 1963 grimly proclaims, "Automation's really here; jobs go scarce—POINT OF NO RETURN FOR EVERYBODY."[82] The decades-old story in photographs is as familiar to a contemporary audience as it is sympathy wrenching. Seasoned machinists and anxious family men describe the existential anguish—and sometimes outright unemployment—that nascent automation introduced to their working lives. In retrospect, we know that 1963 was hardly the "point of no return" for the American worker. Still, the fears wrought by automation of labor seem omnipresent, and sometimes overwhelming.

While some experts worry about the emergence of a perfect substitute for the human worker,[83] others do not see a workless future for mankind. New technologies, they argue, do not always inherently destroy jobs. Although computers may replace humans on assembly lines and in call centers, keeping the new hardware and software systems operational will require humans in support roles.[84] These new job opportunities may be more sophisticated than the old—servicing a robotic arm requires a different skill set than does wielding a power

81. Erik Brynjolfsson and Andrew McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (New York: W. W. Norton and Co., 2014).
82. *LIFE Magazine*, July 19, 1963.
83. Rory Cellan-Jones, "Robots on the March," BBC News, July 2, 2015; Martin Ford, *Rise of the Robots: Technology and the Threat of a Jobless Future* (New York: Basic Books, 2015).
84. David Benady, "Self-Driving Cars to Hospital Robots: Automation Will Change Life and Work," *Guardian*, March 30, 2016.

drill on an assembly line. In the sectors most affected by AI adoption, the nature of work may change fundamentally. Pessimists worry about the inequality generated by disproportionate displacement of workers in the near term.[85] Because AI adoption most readily replaces repetitive, low-skilled labor, jobs that tend to pay less and require lower levels of education face the highest risk of automation.[86] Optimists argue that those who fear mass automation-induced unemployment should consult their history textbooks. Although every technological revolution has rendered certain industries obsolete, history has been punctuated by periods of creative destruction. When one profession becomes unneeded, other job opportunities emerge.[87]

Technological progress has always met resistance and hesitation. So far, the concerns of those who fear a doomsday brought about by laborsaving advancements have generally been unwarranted. More than 2,000 years ago, Aristotle made timorous predictions about technology in his *Politics*.[88] The mass unemployment and societywide idleness that Aristotle feared is not what ultimately led to the demise of Classical Greece. Johannes Trithemius, a monk writing in the late 1400s, decried Gutenberg's printing press as a threat to the monastic way of life.[89] Before printing technology diffused widely, monasteries across Europe harbored scribes who laboriously copied literary works by hand. Pessimists who worried that machines would replace monks were correct—the age of illuminated manuscripts came to an end, but an entirely new economy developed around the printing press. Libraries and newspapers sped the diffusion of knowledge, and writers enjoyed greater connection with audiences. Although the printing press did generate a class of unemployed scribes,

---

85. Aaron Smith and Janna Anderson, "AI, Robotics, and the Future of Jobs" (Pew Research Center, Washington, DC, August 2014).
86. "Today, the main conversation about self-driving cars is not about technological feasibility, but societal impacts and industrial transformation: How difficult will it be for the taxi drivers and truckers who'll lose their jobs to find another way of making a living?" Erik Brynjolfsson, "Technology Is Changing the Way We Live, Learn and Work. How Can Leaders Make Sure We All Prosper?," *World Economic Forum*, January 4, 2017.
87. Ben Miller and Robert D. Atkinson. "Are Robots Taking Our Jobs, or Making Them?" (Information Technology and Innovation Foundation, Washington, DC, 2013), 1.
88. "For if every tool could perform its own work when ordered, or by seeing what to do in advance, like the statues of Daedalus in the story, or the tripods of Hephaestus which the poet says 'enter self-moved the company divine,'—if thus shuttles wove and quills played harps of themselves, master-craftsmen would have no need of assistants and masters no need of slaves." Aristotle, *Politics*, book 1, section 1253b, trans. H. Rackham (London: William Heinemann Ltd., 1932).
89. Redmond A. Burke, "Review of *In Praise of Scribes*, Johannes Trithemius," *Library Quarterly* 45, no. 1 (1975).

the wider world benefited from job opportunities that had not existed before the printing revolution.

Technological progress has accelerated in recent centuries, and such pessimism has recurred more frequently. "There have been periodic warnings in the last two centuries that automation and new technology were going to wipe out large numbers of middle class jobs," observes MIT economist David H. Autor.[90] The industrial revolution of the 18th and 19th centuries faced resistance from critics similar to Aristotle and Johannes Trithemius. Today, we sometimes hear anti-industrialization protestors called "Luddites." Historically, the Luddites were part of a larger antimodernization movement; nearly every early industrial technology met violent resistance.[91] During the last decades of the 18th century and the first of the 19th, "machine breaking" was a frequent and costly response to mechanization.[92] But fears of an apocalypse by automation again were unfounded, and the following two centuries have seen tremendous improvements in the quality of life of the average human. Skeptics often fail to appreciate that while new technologies may obliterate old businesses and jobs, they also allow for many more opportunities that are impossible to foresee.[93] Jobs lost to technological innovation are replaced by work in entirely new sectors that usually offer better wages, a safer work environment, and more leisure time.[94]

In late 2014, economists at Deloitte LLP published a survey of the effect of technology on jobs over the past 200 years. They found that "technology has transformed productivity and living standards, and, in the process, created new

90. David H. Autor, "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives* 29, no. 3 (2015): 3.
91. Alessandro Nuvolari, "The 'Machine Breakers' and the Industrial Revolution," *Journal of European Economic History* 31, no. 2 (2002).
92. Ibid., 146–48.
93. "Discussions of how technology may affect labor demand are often focused on existing jobs, which can offer insights about which occupations may suffer the greatest dislocation, but offer much less insight about the emergence of as-yet-nonexistent occupations of the future." Joel Mokyr, Chris Vickers, and Nicolas L. Ziebarth, "History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?," *Journal of Economic Perspectives* 29, no. 3 (2015): 45.
94. "In the end, the fears of the Luddites that machinery would impoverish workers were not realized, and the main reason is well understood. The mechanization of the early 19th century could only replace a limited number of human activities. At the same time, technological change increased the demand for other types of labor that were complementary to the capital goods embodied in the new technologies. This increased demand for labor included such obvious jobs as mechanics to fix the new machines, but it extended to jobs for supervisors to oversee the new factory system and accountants to manage enterprises operating on an unprecedented scale. More importantly, technological progress also took the form of product innovation, and thus created entirely new sectors for the economy, a development that was essentially missed in the discussions of economists of this time." Ibid., 36.

employment in new sectors."[95] This happens because human needs and tastes change constantly and, therefore, "the stock of work in the economy is not fixed; the last 200 years demonstrates that when a machine replaces a human, the result, paradoxically, is faster growth and, in time, rising employment."[96] It is easy for critics to highlight disruptions in some notable sectors in which machines replaced human labor. The prospects of diminishing employment opportunities or wage stagnation, after all, inflame our most personal and immediate anxieties. Yet the upshot of laborsaving innovations—namely, improved employment quality, expanded labor opportunities, and increased leisure time—receive far fewer media reports and sensationalist book treatments.[97]

The incredible pace of contemporary innovation has brought with it "a resurgence of automation anxiety" in recent years,[98] but the patterns of history generally hold true.[99] Critics will repeat the old argument that *this time it's different!*, but there are good reasons to have faith that humans will once again muddle through and prevail in the face of turbulent, disruptive change. As venture capitalist Marc Andreessen has noted when addressing the fear that automation is running amok and that "robots will eat all the jobs,"

> We have no idea what the fields, industries, businesses, and jobs of the future will be. We just know we will create an enormous number of them. Because if robots and AI [artificial intelligence] replace people for many of the things we do today, the new fields we create will be built on the huge number of people those robots and AI systems made available. To argue that huge numbers of people will be available but we will find nothing for them (us) to do is to dramatically short human creativity. And I am way long [on] human creativity.[100]

---

95. Ian Stewart, Debapratim De, and Alex Cole, "Technology and People: The Great Job-Creating Machine" (London: Deloitte LLP, December 2014), 9.
96. Ibid., 10. The authors note that "machines will take on more repetitive and laborious tasks, but seem no closer to eliminating the need for human labour than at any time in the last 150 years. It is not hard to think of pressing, unmet needs even in the rich world: the care of the elderly and the frail, lifetime education and retraining, health care, physical and mental well-being."
97. Katie Allen, "Technology Has Created More Jobs Than It Has Destroyed, Says 140 Years of Data," *Guardian*, August 14, 2015.
98. Autor, "Why Are There Still So Many Jobs?," 4.
99. "Journalists and even expert commentators tend to overstate the extent of machine substitution for human labor and ignore the strong complementarities between automation and labor that increase productivity, raise earnings, and augment demand for labor." Ibid., 5.
100. Marc Andreessen, "This Is Probably a Good Time to Say That I Don't Believe Robots Will Eat All the Jobs . . . ," *Marc Andreessen* blog, June 13, 2014.

Real-world evidence already supports Andreessen's conclusion that we will learn to adapt to a world full of robots and automated systems. A 2015 economic analysis by Colin Lewis, a behavioral economist and data scientist, showed that "despite the headlines, companies that have installed industrial robots are actually increasingly employing more people whilst at the same time adding more robots." According to Lewis's research, 1.25 million new jobs had been added by companies that make extensive use of industrial robots over the previous six years.[101] This trend held among newer firms such as Amazon and Tesla Motors and among older, more established companies such as Chrysler and Philips Electronics.[102]

It's also worth noting how difficult it is to predict future labor market trends. The jobs that many people worry are disappearing today were, at an earlier point in history, the "disruptors" of an even-older industry. In early 2015, Glassdoor, an online jobs and recruiting site, published a report on the 25 jobs that currently offered the highest pay. Many of these positions would not have made sense to a jobseeker 40 years ago. For example, some desirable jobs on Glassdoor's list[103] included software architect (#3), software development manager (#4), solutions architect (#6), analytics manager (#8), IT manager (#9), data scientist (#15), security engineer (#16), computer hardware engineer (#18), database administrator (#20), UX designer (#21), and software engineer (#23). Looking back at reports from the 1970s and 1980s published by the US Bureau of Labor Statistics, the federal agency that monitors labor market trends, one finds no mention of these information technology–related professions, because they had not yet been envisioned.[104] So what will the most important and well-paying jobs be 30 to 40 years from now? If history is any guide, many of them are beyond what we can imagine.

In the short term, artificially intelligent systems likely will not displace workers in a uniform way across the entire economy. There are several important determinants of how easily a given industry can integrate these technologies into their business models. Scholars at the McKinsey Global Institute argue that capital-intensive industries or those regulated heavily may be insulated from broad AI implementation in the near term.[105] Although a few snippets of code can

101. Colin Lewis, "Study—Robots Are Not Taking Jobs," *Robotenomics*, September 16, 2015.
102. Ibid.
103. Glassdoor, "25 Highest Paying Jobs in Demand," *Glassdoor* blog, February 17, 2015.
104. John Tschetter, "An Evaluation of BLS' Projections of 1980 Industry Employment," *Monthly Labor Review*, August 1984.
105. Michael Chui, James Manyika, and Mehdi Miremadi, "Four Fundamentals of Workplace Automation," *McKinsey Quarterly*, November 2015.

automate a data entry job, the expense and effort required to fully replace an aircraft pilot is much greater; a company interested in fully pilotless airplanes must invest time and money in regulatory compliance, computing research, and hardware upgrades. By comparing the promises of automation technology to the capabilities of present-day workers, we can make educated guesses about who will be outcompeted by robots most quickly. But these predictions become less reliable on the scale of two or three generations—it's hard to confidently say how those yet unborn will earn a paycheck or how business life will be structured in the future.

In a 2013 study, Carl Benedikt Frey and Michael Osborne of the University of Oxford surveyed hundreds of present-day occupations and evaluated the probability that each would face AI-related automation in coming decades.[106] Of the 702 professions analyzed, Frey and Osborne estimated that seamstresses, telemarketers, and library technicians were among the employees who face at least a 99 percent chance of computerization.[107] Among the most future-proof, according to the study, were healthcare providers,[108] artists,[109] and teachers.[110] Aggregately, Frey and Osborne estimate that 47 percent of US jobs are at high risk of automation.[111] It is highly unlikely that 47 percent of US workers will be escorted out of their offices to live out the

"It's hard to confidently say how those yet unborn will earn a paycheck or how business life will be structured in the future."

106. Carl Benedikt Frey and Michael A. Osborne, "The Future of Employment: How Susceptible Are Jobs to Computerization" (Oxford Martin School, University of Oxford, Oxford, UK, 2013).
107. Ibid., 72.
108. In this subset, recreational therapists, mental health counselors, social workers, and surgeons were among professions with less than a 1 percent chance of computerization. These professions require both significant technical skill and "bedside manner." A successful automated replacement, in addition to being highly scientifically sophisticated, would have to exhibit empathy and craft appropriate emotional responses to human patients.
109. Artistic professions least at risk of computerization were choreographers, set designers, fashion designers, and so on. Each of these professions had a greater than 99 percent chance of remaining nonautomated. While recent developments in AI have enabled computer-generated visual arts and music, this is largely still based on vast amounts of existing data.
110. Frey and Osborne, "Future of Employment," 57.
111. Ibid., 47.

rest of their lives unemployed. In every previous case of technological transition, the "Luddite fallacy" has become a tired, widely rejected prophesy of perpetual unemployment. Before productivity-enhancing innovations in agriculture, the vast majority of Americans worked on farms. Today, the agricultural sector employs less than 3 percent of American workers.[112] The livelihoods of millions of Americans were taken over by machines, but those citizens do not sit idle for lack of work in the fields. Today's unemployment figures suggest that increasingly efficient agricultural practices instead allowed many farm laborers to find work in other sectors.

Because labor in some industries is more easily or quickly replaced by computers and automated systems, we expect some degree of asymmetrical transitory unemployment. To remain competitive, those facing unemployment by robot will benefit from remaining flexible and remembering the principle of comparative advantage. Although machines and computers are adept at tooling car parts and delivering relevant results on a search engine site, they are inept at performing in musicals and putting hospital inpatients at ease. The affected workers and businesses will need to adjust to new marketplace realities. That transition will take time. As James Bessen of the Boston University School of Law points out in his book *Learning by Doing*, for technological revolutions to take hold and have a meaningful influence on economic growth and worker conditions, large numbers of ordinary workers must acquire new knowledge and skills. But "that is a slow and difficult process, and history suggests that it often requires social changes supported by accommodating institutions and culture."[113]

Past experience suggests that society will adjust to technological change and standards of living will continue to improve.[114] But according to Bessen, society may not respond immediately to upheavals possible during AI adoption.

The experience of America's textile industry demonstrates that technology sometimes interacts with society in surprising and counterintuitive ways. From the beginning, social experiments have been integral to the introduction of major new technologies, which requires new skills learned through experience. And learning on the job often requires new ways of organizing a workforce, new occupations, and new labor markets.

---

112. "Rethinking the Luddites," *Economist*, November 6, 2009.

113. James Bessen, *Learning by Doing: The Real Connection between Innovation, Wages, and Wealth* (New Haven, CT: Yale University Press, 2015), 223.

114. "But if the past 200 years of innovation have any lesson, it is this: society has repeatedly and quickly integrated and greatly benefited from innovation." Maureen K. Ohlhausen, "Internet of Everything: Data, Networks, and Opportunities" (remarks before the US Chamber of Commerce Foundation, September 22, 2015), 5.

Even with the right policies, these social changes can take time to work out. So while new inventions can come into use relatively quickly, it may take decades of slow learning and occupational changes before the benefits of major new technologies are shared by large numbers of ordinary workers.[115]

In the next few decades, policymakers must carefully consider the possibilities of wage stagnation and rising inequality caused by technological progress. David H. Autor observes that until the 1980s, workers without a college education could expect to earn about 66 percent as much as those with a college degree. The education gap widened over the following decades; by late in the first decade of the 21st century, this ratio had fallen to just over 50 percent. Today's robots artificially increase the supply of labor for jobs that require repetitive actions. In this way, automation displaces workers and exerts downward pressure on wages for easily automated work. As a remedy, Autor suggests that promoting college enrollment is a promising first step toward reducing the disparity between lower- and higher-skilled wages through supply and demand.[116]

Interestingly, data from the second half of the 20th century show that those with the highest and lowest skills have fared better than those in the middle. Maarten Goos and Alan Manning divide the labor market into low-, medium-, and high-skilled work and assert that those on the extremes of the spectrum have jobs that are most difficult to automate, leading to labor polarization. It is not difficult to appreciate why high-skilled workers are comparatively safe, but there is a more subtle difference between the climates for low- and medium-skilled workers. As defined by the authors, the former group tends to occupy itself with manual labor that requires a high degree of physical coordination, like stocking shelves in grocery stores and serving food in restaurants. The latter segment involves itself in assembly lines and data entry—repetitive occupations that are more easily automated.[117]

Because of this difference, societies must emphasize training and preparation for occupations more difficult to computerize—jobs that require creativity, nuanced communication, or compassion. There is already precedent for worker readjustment programs around the world; these initiatives can mitigate trade-induced

---

115. James Bessen, "Will Robots Steal our Jobs? The Humble Loom Suggests Not," *The Switch, Washington Post*, January 25, 2014.
116. David H. Autor, "The Polarization of Job Opportunities in the US Labor Market: Implications for Employment and Earnings" (joint paper, Center for American Progress and the Hamilton Project of the Brookings Institution, Washington, DC, 2010). See also Michael Kiley, "The Supply of Skilled Labour and Skill-Biased Technological Progress," *Economic Journal* 109, no. 458 (1999).
117. Maarten Goos and Alan Manning, "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *Review of Economics and Statistics* 89, no. 1 (2007).

frictional unemployment caused by loss of competitiveness. With the passage of the Trade Expansion Act of 1962, Congress established a system to compensate workers harmed by a more open economy.[118] Whereas trade adjustment assistance in its current form has serious flaws,[119] it may be possible to retool the program and redirect its compensation scheme to those unemployed as a result of technology in an automation adjustment assistance plan. A well-designed initiative could narrow the delay between technological advancement and societal change.

Another potential remedy for technology-related unemployment is a universal basic income, under which every citizen would be guaranteed a certain monthly disbursement from the government. This idea is not new; Milton Friedman expounded the merits of what he called a "negative income tax" in his 1962 book, *Capitalism and Freedom*. Friedman proposed replacing the contemporary welfare apparatus with a simplified system in which lower-income individuals would receive cash transfers on the basis of reported income and other needs previously met by various subsidy programs.[120] On a small scale, Finland[121] and the Netherlands[122] are experimenting with basic income initiatives. Y Combinator, the California startup incubator, has also begun exploring a similar program with families in the San Francisco area.[123]

Current debate on the subject concerns implementation strategy and underlying financial feasibility. Although the idea of a societywide guaranteed income is embraced by a diverse collection of libertarians,[124] conservatives,[125] and progressives,[126] each group has a unique approach. There is disagreement over

118. Trade Expansion Act of 1962, Pub. L. No. 87-794, 76 Stat. 872 (1962). See also "Trade Readjustment Allowances," US Department of Labor, last modified July 10, 2015.
119. David B. Muhlhausen and James Sherk, "Trade Adjustment Assistance, Let the Ineffective and Wasteful 'Job-Training' Program Expire" (Heritage Foundation, Washington, DC, 2014); Sarah Dolfin and Peter Z. Schochet, "The Benefits and Costs of the Trade Adjustment Assistance (TAA) Program under the 2002 Amendments" (final report, Mathematica Policy Research, Princeton, NJ, 2012), 66–69. Dolfin and Schochet offer a caveat: "The negative net benefits [of the trade adjustment assistance (TAA) program] were robust to a wide range of assumptions. However, these calculations do not include the potentially large benefits of the TAA program in making free trade politically feasible."
120. Milton Friedman, *Capitalism and Freedom* (Chicago: University of Chicago Press, 1962).
121. Luke Graham, "Finland Experiments with Universal Basic Income Scheme," *CNBC*, January 3, 2017.
122. Tracy Brown Hamilton, "The Netherlands' Upcoming Money-for-Nothing Experiment," *Atlantic*, June 21, 2016.
123. Jathan Sadowski, "Why Silicon Valley Is Embracing Universal Basic Income," *Guardian*, June 22, 2016.
124. Ed Dolan, "Why Should a Libertarian Take Universal Basic Income Seriously?," *Niskanen Center*, February 6, 2017.
125. David Frum, "A Rule for Conservative Anti-poverty Plans: Keep It Simple," *Atlantic*, July 31, 2014.
126. Scott Santens, "The Progressive Case for Replacing the Welfare State with Basic Income," *TechCrunch*, September 9, 2016; Fred Hiatt, "How Democrats Can Be Progressive without Being Irresponsible," Washington Post, April 9, 2017.

whether a basic income system should supplement[127] or replace[128] existing welfare systems. Such a program would require large annual outlays, and opinions differ widely about how such a large program would be funded. Options include entitlement reform, tax reform,[129] and other, more heterodox approaches like seigniorage.[130]

Alternatively, policymakers could help solve future automation-related wage stagnation by expanding the Earned Income Tax Credit (EITC). At its 2015 level of $67 billion in outlays, the EITC is the third-largest extant welfare effort in the United States, reducing federal tax burdens for 80 percent of those who qualify. According to the Internal Revenue Service and American Community Survey Data, the EITC has removed 9.4 million Americans from poverty.[131] The program was in part a response to suggestions of a negative income tax in the 1960s, as opponents to the latter worried that an unconditionally guaranteed income would reduce incentives to work.[132] Increasing EITC support for low-wage workers could mitigate the immediate problems of rising wage inequality and labor polarization. As Autor indicates, technological progress has pushed medium-skilled workers into lower-skilled jobs rather than rendering them permanently unemployed.[133] The EITC reduces financial pressures on those displaced workers and preserves incentives to stay in the labor market.

Perhaps the solution to the effects of automation in the labor market is not one single government program. An effective scheme for reducing inequality and joblessness could be a bundle of policies, or a combination of public policy and civil society.[134] But eventually, society will become accustomed to shifts in labor markets and the economy as a whole, including more fundamental changes in the goods and services we value.

---

127. Philippe Van Parijs, "A Basic Income for All," *Boston Review* 25, no. 5 (2000).
128. Charles Murray, "Guaranteed Income as a Replacement for the Welfare State," *Basic Income Studies* 3, no. 2 (2008).
129. See Jean-Marie Monnier and Carlo Vercellone, "The Foundations and Funding of Basic Income as Primary Income," *Basic Income Studies* 9, no. 1-2 (2014).
130. See Joseph Huber, "Funding Basic Income by Seigniorage" (paper for the Basic Income European Network 8th Congress, Berlin, October 6–7, 2000). Seigniorage is the practice of financing fiscal policy expenditures through monetary policy. Governments earn revenue from printing new money, but this practice can contribute to runaway inflation.
131. "About EITC," Internal Revenue Service, last modified April 11, 2017, https://www.eitc.irs.gov/EITC-Central/abouteitc.
132. V. Joseph Hotz and John Karl Scholz, "The Earned Income Tax Credit," in *Means-Tested Transfer Programs in the United States*, ed. Robert A. Moffitt (Chicago: University of Chicago Press, 2003).
133. Autor, "Polarization of Job Opportunities in the US Labor Market."
134. Michael Tanner, *The End of Welfare: Fighting Poverty in the Civil Society* (Cato Institute: Washington, DC, 1996).

“We cannot predict with high certainty what labor markets will look like in a decade or two. At present, many jobholders are unprepared for the possible automation of their livelihoods.”

As part of a recent Pew survey of AI experts and futurists, the Monitor Institute’s Tony Siesfeld proposed that small-scale creative goods might become more attractive in a computerized, standardized world.

Entrepreneurially minded unemployed and underemployed people are taking advantage of sites such as Etsy and TaskRabbit to market quintessentially human skills. And in response, demand is increasing for artisanal or handcrafted products that were made by a human.[135]

We cannot predict with high certainty what labor markets will look like in a decade or two. At present, many jobholders are unprepared for the possible automation of their livelihoods. Fortunately, automation will result in net gains to society despite costs to those who find themselves unemployed or displaced to a lower-paying job, and therefore there is a Kaldor-Hicks solution to the robotic threat. Daniel Akst points out,

> Perhaps the biggest lesson we can learn from the midcentury thinkers who worried about automation is that while there is cause for concern, there is no other way but forward. Like trade, automation makes us better off collectively by making some of us worse off. So the focus of our concern should be on those injured by the robots, even if the wounds are “only” economic.[136]

As long as those who stand to lose from automation are appropriately compensated, society must look forward to coming changes. Marc Andreessen optimistically pointed out that we should embrace technological progress and the ever-increasing access to information, markets, and people that it brings.[137] Technology expands opportunity

135. Smith and Anderson, “AI, Robotics, and the Future of Jobs.”
136. Daniel Akst, “What Can We Learn from Past Anxiety over Automation?,” *Wilson Quarterly*, Summer 2013.
137. Andreessen, “This Is Probably a Good Time to Say That I Don’t Believe Robots Will Eat All the Jobs.”

and freedom of action. It is much easier to make a case for technological progress with compensation for those who stand to lose the most than for perpetual resistance to new ideas and potential laborsaving innovation.

Privacy, Discrimination, and Algorithmic Transparency

AI and machine learning have also raised "black box" concerns about the nature of the algorithms and datasets that support these systems. Some commentators argue that the big data and algorithmic matching techniques at the heart of many AI technologies could intentionally or accidentally exacerbate social and economic problems in subtle but powerful ways.

Popularized by law professor Frank Pasquale's book *The Black Box Society: The Secret Algorithms That Control Money and Information*,[138] the term refers to the perceived inscrutability of the processes that result in the digital outputs that users can see. The data that are exchanged when we enter a search term; make online decisions about transactions, employment, housing, and education; or use AI-assisted software programs are framed and filtered by a complex algorithm. Yet such algorithms are often trade secrets, closed from public scrutiny. And even if the public *could* review them, the nature of machine-learning techniques can obviate the usefulness of review because the program is teaching itself. Pasquale and others worry that such opaqueness creates serious problems.

Consider targeted advertisements. When a user enters a query into a search engine, a tailored assortment of advertisements generally appears next to the search results. The advertisements may relate to that specific user's previous online activities or to the object of the search itself. Yet critics allege that the results are anything but objective and can, in fact, be harmful.[139] Latanya Sweeney of Harvard University puts forth this argument in her influential article, "Discrimination in Online Ad Delivery."[140] Sweeney reports that searches for stereotypically African-American names display advertisements suggesting a criminal record, even if that person has no such criminal history. Searches for stereotypically European-American names, on the other hand, display no such criminally suggestive advertisements. Sweeney argues that this kind of disparate racial result can hinder opportunities for employment and housing

138. Pasquale, *Black Box Society*.
139. See Ryan Calo, "Digital Market Manipulation," George Washington Law Review 82 (2014): 999. See also David Talbot, "Data Discrimination Means the Poor May Experience a Different Internet," *MIT Technology Review*, October 9, 2013.
140. Latanya Sweeney, "Discrimination in Online Ad Delivery," *ACMQueue* 11, no. 3 (2013).

for innocent minorities because the individuals may be falsely associated with criminality.

Similar concerns about group stereotyping in algorithmic outputs have been raised about sharing economy platforms,[141] facial recognition software,[142] and web mapping services.[143]

Such concerns are even more pressing when AI software is used in the service of the criminal justice system. Some municipalities have begun to use computer programs that assess the likelihood that a criminal offender will commit another crime in the future. These "risk assessment" scores are then considered by judges who are deciding what kind of sentences to set. Such programs have come under harsh criticism not only for being ineffective in precisely gauging recidivism risk, but also for being biased against African-Americans.[144] A related technology suite known as "predictive policing" uses historical data on crime and demographics to anticipate which neighborhoods should receive aggressive police patrolling. Although police departments claim that such techniques have led to crime reductions in targeted areas, critics allege that predictive policing is little more than a fishing expedition for largely nonviolent offenders.[145]

Nor are stereotypes the only concern. Large internet companies have long been criticized for their potential to abuse market power in an anticompetitive or corrupt manner that is at odds with the public interest. For example, critics allege that a social network could decide to manipulate algorithms to promote news that puts the company in a favorable light or downplay negative news.[146] Or a technology company could be pressured by a government to censor certain information or promote state propaganda.[147]

Even without those concerns about possible discrimination or corruption, yet another group of critics scrutinize AI technologies for the perceived privacy problems that they could pose.[148] Neil M. Richards and Jonathan H. King,

141. See, for example, Benjamin G. Edelman and Michael Luca, "Digital Discrimination: The Case of Airbnb.com" (Working Paper Number 14-054, Negotiations, Organizations, and Markets Unit, Harvard Business School, Cambridge, MA, 2014).
142. Clare Garvie and Jonathan Frankle, "Facial-Recognition Software Might Have a Racial Bias Problem," *Atlantic*, April 7, 2016.
143. Andrew Marantz, "When an App Is Called Racist," *New Yorker*, July 29, 2015.
144. Julie Angwin et al., "Machine Bias," *ProPublica*, May 23, 2016.
145. Maurice Chammah, "Policing the Future," *Marshall Project*, February 3, 2016.
146. Michael Nunez, "Former Facebook Workers: We Routinely Suppressed Conservative News," *Gizmodo*, May 9, 2016.
147. Jordan Pearson, "Is Twitter Censoring a Blockbuster Report on US Drone Assassinations?," *Motherboard*, October 19, 2015.
148. Ryan Calo, "Peeping HALs: Making Sense of Artificial Intelligence and Privacy," *European Journal of Legal Studies* 2, no. 3 (2010).

for example, argue that big data algorithmic techniques, by their very nature, could spell "the end of privacy" in the quest to bring insight into online behaviors.[149] Worse, while such techniques can ostensibly make the actions of millions of online users more transparent to those wielding them, they are shrouded in mystery to the average user. To overcome this so-called transparency paradox, the authors advocate for new forms of hard and soft regulation that would open big data techniques to government audit or punitive actions.[150]

Other critics are worried about the ethical implications of autonomous vehicles[151] and specifically the decisions that algorithms powering autonomous systems are programmed to make in life-and-death situations.[152]

Many technology companies, for their part, tend to share the concerns about false or damaging algorithmic outputs. (In fact, Google was one of the funders of Sweeney's research on discrimination and advertisements.) After all, these companies make money in part by delivering relevant content to the right parties. Firms have a very real profit incentive to improve their algorithms so that they are as accurate and useful as possible. And if any one company does not do so, "our competition expertise tells us that if one company draws incorrect conclusions and misses opportunities, competitors with better analysis will strive to fill the gap," in the words of Federal Trade Commission (FTC) Commissioner Maureen K. Ohlhausen.[153] In the case of driverless cars and drones, the difference between a perfect and near-perfect autonomous system can spell the difference between life and death. And on a personal level, many of the engineers who work on such algorithms do not wish to develop a reputation for furthering the social ills of racism, sexism, or corruption.

To the extent that these conversations have been collaborative and constructive, they have served to assist firms as they strive to constantly improve

---

149. Neil M. Richards and Jonathan H. King, "Three Paradoxes of Big Data," *Stanford Law Review* 66, no. 41 (2013).
150. Neil M. Richards and Jonathan H. King, "Big Data and the Future for Privacy," in *Research Handbook on Digital Transpormations*, ed. F. Xavier Olleros and Majlinda Zhegu (Cheltenham, UK: Edward Elgar, 2016).
151. Patrick Lin, "The Ethics of Saving Lives with Autonomous Cars Is Far Murkier Than You Think," *Wired*, July 30, 2013.
152. Keith Kirkpatrick, "The Moral Challenges of Driverless Cars," *Communications of the ACM* 58, no. 8 (2015). For a response to those concerns, see Adam Thierer, "On the Line between Technology Ethics vs. Technology Policy," *Technology Liberation Front*, August 1, 2013; Adam Thierer, "Making Sure the 'Trolley Problem' Doesn't Derail Life-Saving Innovation," *Technology Liberation Front*, January 13, 2015.
153. Maureen K. Ohlhausen, "Appendix: Separate Statement of Commissioner Maureen K. Ohlhausen," in Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion?* (FTC staff report, January 2016).

their products so that they are useful as possible. Unfortunately, some concerns have metastasized into unproductive and potentially harmful calls for excess government control of technological developments in AI and even mandates of so-called algorithmic transparency.[154] Such policies could have the adverse effect of requiring innovators to first ask for permission from government bureaucrats before improving their products or could place firms in an uncertain environment in which works-in-progress could be audited or halted by possibly uninformed commissars from above.

For example, a 2014 report from the White House Office of Science and Technology Policy encourages government agencies and offices such as the Justice Department, Consumer Financial Protection Bureau, and Equal Employment Opportunity Commission to "identify practices and outcomes facilitated by big data analytics that have a discriminatory impact on protected classes, and develop a plan for investigating and resolving violations of law."[155] The report also hints at possible new economic regulations, calling on the Council of Economic Advisers to "assess the evolving practices of differential pricing" to "consider whether new practices are needed to ensure fairness for the consumer." If not carefully tailored to cover clear and precise violations of existing law, this kind of new regulatory authority could have the potential to impose a kind of chilling effect on AI development. Few innovators will want to undertake projects that carry an extreme risk of running afoul of some ill-defined and potentially retaliatory regulatory regime.

Similarly, the FTC has shown interest in launching enforcement actions against big data applications. After discussing some of the promises and challenges associated with big data technologies that we have described, the report considers how the FTC should treat these technologies. While noting that big data outcomes may have disparate impacts on certain groups "even when data analysts are very careful" in designing their programs, the report nonetheless states that the FTC will bring enforcement actions in areas in which "big data practices *could* violate existing laws" (emphasis added).[156] Ohlhausen, in her separate statement, correctly points out the risk that an improper application of these authorities could discourage "the development of the very tools that promise new benefits to low income, disadvantaged, and vulnerable individuals."[157]

---

154. Katherine Noyes, "The FTC is Worried about Algorithmic Transparency, and You Should Be Too," *PCWorld*, April 9, 2015.
155. Podesta et al., *Big Data: Seizing Opportunities, Preserving Values*.
156. Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion?*
157. Ibid.

Indeed, the 2016 White House report on big data technologies offers four examples of applications that can help alleviate traditional barriers for low-income groups to gain access to credit, jobs, education, and justice.[158]

Others in academia and the media call for even more onerous regulations. Law professor Danielle Keats Citron advocates for a "technological due process," which she describes as a "carefully structured inquisitorial model of quality control" over algorithms and AI technologies.[159] "Inquisitorial" is a choice word: Citron envisions an FTC that would be empowered to extract and audit proprietary software programs at its discretion.[160] Pasquale goes even farther, advocating for the creation of an entirely new body called the Federal Search Commission that would operate like the FCC. He argues that "some governmental agent should be able to peer into the black box of search and determine whether or not illegitimate manipulation has occurred."[161]

Such proposals unfortunately raise more questions than they answer. Importantly, the very nature of machine learning necessarily limits the effectiveness of even the most ideal implementation of these ideas: it may be impossible to scrutinize the outcomes of these programs by analyzing their code because the programs are intended to learn on their own. Furthermore, critics who suggest extending the authority of existing regulatory bodies or the creation of a new regulatory body to audit algorithms, datasets, and techniques actually advance a "transparency paradox" of their own.

On the one hand, such critics recognize the commercial value of such trade secrets, and they generally stop short of calling for a mandated out-and-out publication of all or most algorithms. On the other hand, they maintain that people and businesses can trust faraway bureaucrats to fairly and accurately scrutinize such programs on their behalf. This is not transparency at all, but rather the addition of a new opaque authority with potentially extensive power over a promising new industry. Rather than promoting openness, such a body could counterproductively become captured by interests that are actually antithetical to the public interest, thereby creating new problems while ignoring or even exacerbating old ones.

158. Cecilia Muñoz, Megan Smith, and DJ Patil, *Big Data: Algorithmic Systems, Opportunity, and Civil Rights*.
159. Danielle Keats Citron, "Technological Due Process," *Washington University Law Review* 85 (2007).
160. Danielle Keats Citron, "Big Data Should Be Regulated by 'Technological Due Process,'" *Room for Debate*, *New York Times*, July 29, 2016.
161. Frank Pasquale, "Internet Nondiscrimination Principles for Competition Policy Online" (Testimony before the Task Force on Competition Policy and Antitrust Laws of the House Committee on the Judiciary, July 15, 2008).

There is a middle ground. Just as technology companies are not eager to gain reputations as mechanisms for furthering social ills, they would also be harmed if their services developed reputations as out-of-control algorithmic deities with no scrutiny or accountability. Indeed, several large technology firms have released some major AI products as "open source" software, which means that the code is fully available for the public to view and even contribute to.[162] Regulators or nongovernmental oversight bodies could encourage businesses to share more details of AI applications as appropriate. There are even technological tools that could be employed to demonstrate procedural compliance or exonerate allegations of discrimination without fully divulging a technology's source code.[163]

We should not, however, expect that businesses will always voluntarily divulge the specific information that critics desire. For many reasons, we might not want them to.[164] For example, publicizing the contours of a particular AI application may lead opportunistic parties to game the service for their own benefit, thereby undermining the service's effectiveness.[165] Developers spend a lot of time testing and tweaking their applications to discern exactly what the effects of different changes will be. There is a lot of trial and error involved. This is not to say that any given application will quickly become as close to perfect as possible, but rather that the solutions that may look obvious from afar can engender many ill consequences that are not immediately obvious to either the regulator or even the creator.

As with any new technology, the rise of artificial intelligence techniques has brought with it a score of criticisms and concerns. Some of these are overblown or minor. When such small issues do arise—as when a search engine or smart assistant provides inaccurate or unhelpful results, for example—they are generally quickly and adequately corrected by the appropriate party. Such applications of AI technologies are clearly distinct from those that may pose risks of human casualties or government prosecution, such as in the case of predictive policing. If an unnecessary new kind of administrative regulation or liability regime meant for the second category of applications were applied to the first,

---

162. Cade Metz, "Google Just Open Sourced TensorFlow, Its Artificial Intelligence Engine," *Wired*, November 9, 2015.
163. Joshua A. Kroll et al., "Accountable Algorithms," *University of Pennsylvania Law Review* 165 (forthcoming, 2017).
164. Mike Ananny and Kate Crawford, "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability," *New Media & Society* (December 2016).
165. Nicholas Diakopolous, "Accountability in Algorithmic Decision Making," *Communications of the ACM* 59, no. 2 (2016).

largely self-correcting kinds of applications, the damage to innovation could be significant. In addition, the same disadvantaged communities that critics seek to assist would also be robbed of the benefits that such technologies could bring.

Other concerns, by nature of their integration into government administration and criminal justice, are more pressing. In particular, so-called predictive policing and criminal sentencing software suites can create unique civil liberties concerns for their direct effect on the administration of justice. In such circumstances, more public oversight and accountability mechanisms are perhaps appropriate. Yet this scrutiny would merely be an extension of the same transparency that we expect of all government contractors. It is the relationship between these programs and the state, and not the nature of the programs themselves, which warrants such extra scrutiny.

Of course, "any fair assessment of algorithms must be made against their alternative," notes Anupam Chander of the School of Law at the University of California, Davis. Specifically, they must be judged against *human* failings that might be better brought to light through machine learning because "these systems will reflect human intention," which may be inherently biased to begin with.[166] As Chander elaborates,

> Algorithms are certainly obscure and mysterious, but often no more so than the committees or individuals they replace. The ultimate black box is the human mind. . . . The consciously racist or sexist algorithm is less likely than the consciously or unconsciously racist or sexist human decision-maker it replaces.[167]

This is not to say that government has no role in overseeing AI technologies. But it may be the case that many of the concerns related to privacy, bias, and discrimination are

"It may be the case that many of the concerns related to privacy, bias, and discrimination are already covered by existing laws and regulations that address human failings in this regard."

166. Atkinson, "It's Going to Kill Us!," 26.
167. Anupam Chander, "The Racist Algorithm?," *Michigan Law Review* 115 (2017): 1030.

already covered by existing laws and regulations that address *human* failings in this regard. Or there may exist other, less restrictive ways of dealing with those perceived problems.

Problems with Precautionary Regulation

It is important to provide a more detailed explanation of the problems associated with traditional regulatory proposals and processes that stand on precautionary principle–based reasoning. Traditional administrative regulatory systems tend to be overly rigid, bureaucratic, inflexible, and slow to adapt to new realities. This approach is particularly problematic as it pertains to the governance of new, fast-moving technologies.

Prior restraints on innovative activities are a recipe for economic and social stagnation. By focusing on preemptive remedies that aim to predict hypothetical problems that may not ever come about, regulators run the risk of making bad bets that are based on their inability to predict the future. Attempting to preemptively plan for every hypothetical worst-case scenario and then require it to be addressed through a regulatory process means that many *best-case* scenarios will never come about.[168] Worse yet, adopting prior restraints on AI experimentation could also undermine efforts to discover bottom-up solutions to many of the hard problems that some critics want addressed preemptively. Oftentimes, it is only through ongoing trial and error that we discover sensible solutions to the legitimately difficult challenges that new technologies can pose: "Both individuals and institutions learn how to do things better—both more efficiently and safer—by making mistakes and dealing with adversity."[169]

The late political scientist Aaron Wildavsky thoroughly documented the deficiencies of anticipatory regulatory efforts that were rooted in precautionary principle–based reasoning:

> Regulation, because it deals with the general rather than with the particular, necessarily results in forbidding some actions that might be beneficial. Regulators cannot devise specifications sufficiently broad to serve as guidelines for every contingency without also limiting some actions that might increase safety. Because regulation is anticipatory, regulators frequently guess

168. Thierer, *Permissionless Innovation*, 2.
169. Adam Thierer, "Failing Better: What We Learn by Confronting Risk and Uncertainty," in *Nudge Theory in Action: Behavioral Design in Policy and Markets*, ed. Sherzod Abdukadirov (London: Palgrave Macmillan, 2016).

> wrong about which things are dangerous; therefore, they compensate by blanket prohibitions.[170]

This risk is perhaps more pronounced when dealing with AI technologies.[171] As mentioned earlier, *how* artificial intelligence is regulated makes little sense until policymakers define *what* it actually entails. As noted, the boundaries of AI are amorphous and ever changing. Artificial intelligence technologies are already all around us—examples include voice-recognition software, automated fraud detection systems, and medical diagnostic technologies—and new systems are constantly emerging and evolving rapidly.[172] And indeed, the nature of AI technologies may frustrate attempts at ex ante regulation altogether.[173]

Policymakers should keep in mind the rich and distinct variety of opportunities presented by artificial intelligence technologies, lest regulations more appropriate for one kind of application inadvertently stymie the development of another and lead to unintended consequences.

For example, a recent filing to the National Highway Traffic Safety Administration (NHTSA) from the Mercatus Center at George Mason University documented the potential opportunity costs that might result from the efforts to pigeonhole emerging autonomous vehicles systems into traditional regulatory processes.[174] The report found that the longer the delay for implementing driverless car technology, the greater the cost to human life. Thousands of individuals are harmed or killed on the road each year because of driver error or intoxication. Delaying driverless car technology, which obviates the risk of such casualties, therefore needlessly increases the number of highway deaths. For example, if regulatory backlog slows the deployment of driverless cars by 5 percent, the authors project an additional 15,500 fatalities over the course of 31 years. A 10 percent delay implies an additional 34,600 fatalities over 33 years,

---

170. Aaron Wildavsky, *Searching for Safety: Social Theory and Social Policy* (New Brunswick, CT: Transaction Books, 1988), 183.
171. Scherer, "Regulating Artificial Intelligence Systems."
172. AJ Agrawal, "7 Ways Artificial Intelligence is Improving Consumer Experiences," *Customer Think*, July 14, 2016.
173. "Ex ante regulation would be difficult because AI research and development may be discreet (requiring little physical infrastructure), discrete (different components of an AI system may be designed without conscious coordination), diffuse (dozens of individuals in widely dispersed geographic locations can participate in an AI project), and opaque (outside observers may not be able to detect potentially harmful features of an AI system)." Scherer, "Regulating Artificial Intelligence Systems," 356–57.
174. Adam Thierer and Caleb Watney, "Comment on the Federal Automated Vehicles Policy Docket" (Public Interest Comment, Mercatus Center at George Mason University, Arlington, VA, November 22, 2016).

whereas a 25 percent delay is projected to cause an additional 112,400 fatalities over 40 years. The case of driverless cars presents a salient example of how an overabundance of caution can actually do much more harm than good.

"If we want progress—an increase in economic growth, improved health, a better environment, etc.—then it is time to regain our sense of optimism about the promise of technological innovation," argues Robert Atkinson of the Information Technology and Innovation Foundation. "In particular, when it comes to AI, we should be enthusiastic and excited, not fearful and cautious."[175]

Our next section will outline an alternative approach to automated and big data technologies that will assuage the critics' biggest concerns while protecting spaces for improvement and innovation. Collaboration, not control, is key.

WHAT THE INTERNET TEACHES US ABOUT TECHNOLOGY GOVERNANCE

As policymakers consider the governance of AI, they would be wise to consider the lessons that can be drawn from America's recent experience with the internet and public policy for the digital economy. Such an examination is particularly relevant here because most AI technologies tap the same building blocks that power the digital economy: code, computers, massive databases and storage capacity, tracking and geolocation technologies, and multilevel dialogue among developers, industry, academics, and government representatives.

Today these technologies are so ubiquitously integrated into almost every facet of our lives and economy that it is easy to take them for granted. Just 20 years ago, however, most of those digital technologies did not yet exist, and no one could have predicted the explosion of new companies and choices that was about to occur.

The United States made several crucial public policy decisions in the early to mid-1990s that helped bring about this momentous technological revolution. Early, explicit guidance from the US government secured a space for innovation and commercial activity on the internet. In contrast, the European Union (EU) pursued a much more prohibitive set of policies intended to protect data privacy. We will discuss each contrasting approach and the according outcomes in the following sections.

175. Atkinson, "It's Going to Kill Us!," 10.

The American Experience: Permissionless Innovation as Policy

In 1994, the Clinton administration decided to allow open commercialization of what was previously just the domain of government agencies and university researchers. Shortly thereafter, Congress passed and President Bill Clinton signed the Telecommunications Act of 1996, which avoided regulating the internet like analog-era communications and media technologies. Notably, section 230 of that legislation explicitly exempted online intermediaries from bearing liability for content that users might post on their platforms.[176]

More importantly, in 1997 the Clinton administration released its *Framework for Global Electronic Commerce*, which articulated the US government's approach toward the internet and the emerging digital economy.[177] The framework was a succinct, market-oriented vision for cyberspace governance that recommended reliance on civil society, contractual negotiations, voluntary agreements, and ongoing marketplace experiments to solve information-age problems.[178] Specifically, the framework recommended that "the private sector should lead [and] the Internet should develop as a market driven arena not a regulated industry,"[179] and that governments should "avoid undue restrictions on electronic commerce."[180] And, finally, that "where governmental involvement is needed, its aim should be to support and enforce a predictable, minimalist, consistent and simple legal environment for commerce."[181]

The combined effect of these policy pronouncements was to encourage a culture of permissionless innovation in that innovators were generally left free to experiment with new technologies and business models.[182]

America's embrace of this policy model helped propel the rise of e-commerce, online speech, and the modern digital revolution.[183] It also helped catapult US-based tech firms into a dominant global position and made them familiar names across the world.[184]

---

176. Derek Khanna, "The Law That Gave Us the Modern Internet—and the Campaign to Kill It," *Atlantic*, September 12, 2013.
177. White House, *The Framework for Global Electronic Commerce*, July 1997.
178. Adam Thierer, "15 Years On, President Clinton's 5 Principles for Internet Policy Remain the Perfect Paradigm," *Forbes*, February 12, 2012.
179. White House, *Framework for Global Electronic Commerce*.
180. Ibid.
181. Ibid.
182. Vinton Cerf, "Keep the Internet Open," *New York Times*, May 24, 2012.
183. Adam Thierer, "Embracing a Culture of Permissionless Innovation," *Cato Online Forum*, November 2014.
184. Adam Thierer, "How Attitudes about Risk and Failure Affect Innovation on Either Side of the Atlantic," *PlainText*, June 19, 2015; Stephen Ezell and Philipp Marxgut, "Comparing American and

To highlight the benefits of the permissionless innovation approach to technology policy, the Mercatus Center has recently published a book,[185] a series of law review articles, and several agency filings that explain what a permissionless innovation policy vision would entail for many different technologies and sectors, including the internet of things and wearable devices,[186] smart cars,[187] commercial drones,[188] Bitcoin,[189] 3-D printing,[190] robotics,[191] the sharing economy,[192] and advanced medical devices and applications.[193]

As identified in those studies, the three key attributes of a permissionless innovation approach to technology policy are as follows:

European Innovation Cultures," in *Shaping the Future: Economic, Social, and Political Dimensions of Innovation* (Austrian Council for Research and Technology Development, 2015), 193. Ezell and Marxgut explain, "Cultural aspects have a significant impact on innovation and inform how entrepreneurial countries, organizations, and people can be. The United States maintains the world's most vibrant innovation culture, where risk and failure are broadly tolerated, inquiry and discussion are encouraged, and the government's role in business plays a less prominent role. . . . There are elements in the European innovation culture that need improvement: a simpler regulatory environment, a broader availability of risk capital, and more tolerance of risk and change being critically important."

185. Thierer, *Permissionless Innovation*.

186. Adam Thierer, "The Internet of Things and Wearable Technology: Addressing Privacy and Security Concerns without Derailing Innovation," *Richmond Journal of Law and Technology* 21, no. 6 (2015).

187. Adam Thierer and Ryan Hagemann, "Removing Roadblocks to Intelligent Vehicles and Driverless Cars," *Wake Forest Journal of Law and Policy* 5, no. 2 (2015).

188. Jerry Brito, Eli Dourado, and Adam Thierer, "Federal Aviation Administration: Unmanned Aircraft System Test Site Program" (Public Interest Comment, Mercatus Center at George Mason University, Arlington, VA, April 23, 2013); Eli Dourado, "The Next Internet-Like Platform for Innovation? Airspace. (Think Drones)," *Wired*, April 23, 2013; Adam Thierer, "Filing to FAA on Drones and 'Model Aircraft,'" *Technology Liberation Front*, September 23, 2014.

189. Jerry Brito and Andrea Castillo O'Sullivan, *Bitcoin: A Primer for Policymakers*, 2nd ed. (Arlington, VA: Mercatus Center at George Mason University, 2016).

190. Adam Thierer and Adam Marcus, "Guns, Limbs, and Toys: What Future for 3D Printing?," *Minnesota Journal of Law, Science, and Technology* 17, no. 2 (2016).

191. Adam Thierer, "Problems with Precautionary Principle–Minded Tech Regulation and a Federal Robotics Commission," *Medium*, September 22, 2014.

192. Christopher Koopman, Matthew Mitchell, and Adam Thierer, "The Sharing Economy and Consumer Protection Regulation: The Case for Policy Change," *Journal of Business, Entrepreneurship, and the Law* 8, no. 2 (2015); Adam Thierer et al., "How the Internet, the Sharing Economy, and Reputational Feedback Mechanisms Solve the 'Lemons Problem,'" *University of Miami Law Review* 70 (2016).

193. Richard Williams, Robert Graboyes, and Adam Thierer, "US Medical Devices: Choices and Consequences" (Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, 2015); Adam Thierer, "The Right to Try and the Future of the FDA in the Age of Personalized Medicine" (Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, 2016).

- *Avoid prior restraints*. Constraints on new innovation should be the last resort, not the first. Innovation should be innocent until proven guilty. Just as importantly, policymakers should make this default position of permissionless innovation clear to would-be entrepreneurs. Uncertainty over the future of regulation can be just as damaging to innovation as limited, but well-understood, regulation. Businesses should not fear that new ventures would be retroactively punished for violating some arcane and unsuitable regulation long after the fact. When appropriate, policymakers should quickly clarify the extent to which previous rules apply to new technologies. When old regulations would unduly quash the innovative potential of a new technology, policymakers should quickly promulgate liberalizing reforms.
- *Base policy on evidence, not fear*. Policymakers should not base policy on worst-case hypotheticals. If public policy is guided at every turn by fear of hypothetical worst-case scenarios and the precautionary mindset, then innovation becomes less likely. Opponents of technological change should bear the burden of proving the harms they allege. And problems that develop are usually best addressed in an ex post fashion.[194]
- *Flexible, bottom-up solutions are better than rigid, top-down controls*. The best solutions to complex social problems are almost always organic and bottom-up in nature. Education and empowerment, social pressure, societal norms, voluntary self-regulation, and targeted enforcement of existing legal norms (especially through the common law) are almost always superior to top-down, command-and-control regulatory edicts and bureaucratic schemes of a permissioned nature.

In practical terms, the problem with highly precautionary regulation is that it results in fewer services, lower-quality goods, higher prices, diminished economic growth, and a decline in the overall standard of living.[195] When public policy is shaped by precautionary principle reasoning, it poses a serious threat to technological progress, economic entrepreneurialism, social adaptation, and long-run prosperity.

194. Adam Thierer, "What 20 Years of Internet Law Teaches Us about Innovation Policy," *Federalist Society* blog, May 12, 2016; Ithiel de Sola Pool, *Technologies of Freedom: On Free Speech in an Electronic Age* (Cambridge, MA: Belknap Press of Harvard University Press, 1983), 231. Regarding regulation of information markets, Pool stresses that "enforcement must be after the fact, not by prior restraint" and that "regulation is a last recourse. In a free society, the burden of proof is for the least possible regulation of communication."
195. Thierer, *Permissionless Innovation*.

To the maximum extent possible, the default position toward new forms of technological innovation should be "innovation allowed." The burden of proof rests on those who favor precautionary regulation to explain why ongoing experimentation with new ways of doing things should be prevented preemptively.

"European countries clearly have the potential to be innovative. . . . Why, then, does Europe still generally lag?"

Europe's Alternative Policy Framework: The Precautionary Principle Approach

While the US digital technology sector prospered in a permissionless innovation policy environment, European policymakers adopted a very different policy framework for the continent's digital technology sector.

The outcomes from the approaches that each rival took to technology policy are stark and instructive. The United States leads the world in technological innovation and houses some of the most well-known and successful technology firms. In 2015, 11 of the top 20 market-leading internet companies were based in the United States. The remaining 9 were in China, Japan, and the Republic of Korea.[196] Noticeably absent was the presence of any firm hailing from the EU, given the EU's wealth of human and productive capital.

European countries clearly have the potential to be innovative. They often boast highly educated workforces, impressive living standards, ample investment capital, and high levels of trade with the countries that do host innovative firms. Why, then, does Europe still generally lag?

This question is one that European policymakers and economists, understandably, are particularly interested in answering. A growing consensus is emerging that a large part of the divergence in outcomes is fundamentally cultural.[197] The United States in general, and Silicon Valley in particular,

196. Mary Meeker, "Internet Trends 2015" (Kleiner Perkins Caufield Byers, Menlo Park, CA, May 27, 2015).
197. Ezell and Marxgut, "Comparing American and European Innovation Cultures."

is thought to embrace risk and the lessons learned from failure. In Europe, on the other hand, "failure is regarded as a personal tragedy," according to German-born economist Petra Moser. Indeed, the European Commission produced a report examining the large gap in innovation that exists between the United States and the EU.[198] It noted that the United States exhibits a much more entrepreneurial culture that embraces risk-taking and failure, whereas European cultures tend to avoid risk and minimize the chance of failure.[199] Many European policies, accordingly, have the effect of openly discouraging risk-taking and entrepreneurship.[200]

The EU's approach is rooted in the precautionary principle. Precautionary principle reasoning refers to the belief that new innovations should be curtailed or disallowed until their developers can demonstrate that they will not cause any harms to individuals, groups, specific entities, cultural norms, or various existing laws or traditions.[201]

In the European approach to technology, fears about privacy and data fueled a series of restrictive policies in the 1990s and beyond that have served to inadvertently stifle the innovative potential of the EU.

The 1995 EU Data Protection Directive, for example, instituted a relatively restrictive set of regulations for online data collection and use.[202] This policy served to largely stymie the development of targeted ad–based business models that propelled US companies such as Google and Facebook to great success.[203] Indeed, the United States did not pursue such heavy-handed privacy regulations as the EU did, thereby increasing the likelihood of innovation in the United States while it diminished in the EU.[204] This difference likely had a cumulative effect. Firms in EU countries struggled to secure venture capital in a regulatory environment that discouraged innovation.[205] With less capital available, firms

198. Simon Forge et al., "Comparing Innovation Performance in the EU and the USA: Lessons from Three ICT Sub-sectors" (JRC Technical Report No. EUR 25961 EN, Institute for Prospective Technological Studies, Joint Research Council, European Commission, Seville, Spain, 2013).
199. Ibid., 46–48.
200. James B. Stewart, "A Fearless Culture Fuels U.S. Tech Giants," *New York Times*, June 18, 2015.
201. See, for example, Roberto Andorno, "The Precautionary Principle: A New Legal Standard for a Technological Age," *Journal of International Biotechnology Law* 1, no. 1 (2004).
202. Directive 95/46/EC of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, October 1995.
203. Avi Goldfarb and Catherine Tucker, "Privacy and Innovation," in *Innovation Policy and the Economy*, vol. 12, ed. Josh Lerner and Scott Stern (Chicago: University of Chicago Press, 2012).
204. Tal Z. Zarsky, "The Privacy-Innovation Conundrum," *Lewis and Clark Law Review* 19, no. 1 (2015).
205. Josh Lerner, "The Impact of Privacy Policy Changes on Venture Capital Investment in Online Advertising Companies" (White Paper for the Analysis Group, Menlo Park, CA, 2012).

were less equipped to discover the kinds of game-changers that characterized the most successful "unicorn" firms in the United States.[206]

The difference in outcomes evident between the EU and United States is remarkable. On one hand, the foresight and leadership of the Clinton administration set the tone and vision in the United States that embraced risk-taking, innovation, and collaboration. The EU, on the other hand, allowed worst-case thinking and heavy-handed proactive regulations to bog down a burgeoning industry before it even had a chance to develop. Importantly, many of these European data regulations could come to encumber AI systems in the near future.[207]

Implications for AI Policy

If policymakers wish to replicate the success we have seen over the past 20 years with the internet, they need to adopt a similar light-touch approach for the governance of AI systems and technologies. A permissionless innovation policy approach will be essential if we hope to capture the profound potential benefits associated with AI technologies. Just as the Clinton administration's embrace of permissionless innovation helped to spur the digital revolution, policymakers can extend that ethos to new sectors such as AI to help fuel similar technological revolutions.[208]

Policymakers who consider how to approach AI technologies today stand in a position similar to that of policymakers who observed the rise of internet activities in the 1990s. Developments that have been hotly anticipated by both professional scientific communities and the heady world of science fiction are now being applied to a range of activities on a continuum of intensity. Lighthearted social media activities are being enhanced by autonomous recommendation services and digital assistants. Medical interventions can be partially or

206. A "unicorn" firm is a startup valued at over $1 billion. Because of the rarity of these kinds of ventures, they are compared to the mythical beast. Samuel Kortum and Josh Lerner, "Assessing the Contribution of Venture Capital to Innovation," *RAND Journal of Economics* 31, no. 4 (2000).
207. Cade Metz, "Artificial Intelligence Is Setting Up the Internet for a Huge Clash with Europe," *Wired*, July 11, 2016.
208. Good examples of how legislators can promote permissionless innovation in their policy pronouncements can be found in some of the speeches of senators Cory Booker (D-NJ) and Deb Fischer (R-NE), and former Senator Kelly Ayotte (R-NH). See Adam Thierer, "A Nonpartisan Policy Vision for the Internet of Things," *Technology Liberation Front*, December 11, 2014; Adam Thierer, "What Cory Booker Gets about Innovation Policy," *Technology Liberation Front*, February 16, 2015. Similarly, no regulator in recent memory has done more to promote permissionless innovation as a policy guidepost than Maureen K. Ohlhausen, a commissioner with the Federal Trade Commission. See Adam Thierer, "FTC's Ohlhausen on Innovation, Prosperity, 'Rational Optimism' and Wise Tech Policy," *Technology Liberation Front*, September 25, 2015.

completely guided with the help of AI technologies. The challenges posed by different kinds of applications of AI technologies are as diverse as the appropriate kinds of solutions that will address public concerns without harming the potential for innovation.

First, policymakers should take care to understand and distinguish the myriad applications of AI technologies so that they are well equipped to appropriately address each kind. Regulations intended to address experimental medical applications of AI technologies, for instance, should not inadvertently be applied to benign social media applications because of broad or improper wording. Some types of AI technologies could be exempt from new regulations entirely, whereas others that pose more direct risks to safety or health could be studied to determine what kind of oversight is appropriate. Properly tailoring the right policies to the right technologies takes time and humility, but it will be well worth the effort.

Next, policymakers should consider risks and concerns in a rational, productive manner rather than allowing worst-case scenario thinking to guide decisions. Policymakers have a general tendency to cleave to an excessively risk-averse approach to regulations. This precautionary-principle threat is perhaps more pronounced in the case of AI technologies, which have been chronicled for better or for worse as a god-like invention spawning terrific and terrible outcomes for the humans who invoke them.[209] But life is not a science fiction movie. Policymakers should take care to separate fantasy from reality when addressing automated technologies.

Finally, policymakers should, in general, embrace a vision of permissionless innovation so that we can enjoy as many benefits from innovation as possible. The lesson of the rivalrous policy approaches to the internet in the 1990s is instructive. Policymakers can either follow the example of the United States and prioritize a clear space for experimentation and commercialization that engenders collaboration and growth, or they can follow the path of the EU and inadvertently quash an industry before it has the chance to develop.

Patience and a general openness to permissionless innovation represent the wise disposition toward new AI technologies not only because it provides breathing space for future entrepreneurialism and invention, but also because it provides an opportunity to see how societal attitudes toward new technologies

209. "When it comes to AI, policymakers should rely on the innovation principle, not the precautionary principle. In other words, we should proceed on the assumption that AI will be fundamentally good, and while it will present some risks, as every technology does, we should focus on addressing these risks, rather than slowing or stopping the technology." Robert D. Atkinson, "It's Going to Kill Us!"

evolve. As the adage goes, "nothing ventured, nothing gained." We will now discuss a few of the benefits that AI technologies may deliver, provided that we get our policies right.

THE CONSTRUCTIVE PATH FORWARD: COLLABORATION, NOT CONTROL

We suggest that a different policy approach for AI is needed, one that is rooted in humility and a recognition that we possess limited knowledge about the future. Policymakers would be wise to heed the advice of FTC Commissioner Ohlhausen, who suggested that

> It is . . . vital that government officials, like myself, approach new technologies with a dose of regulatory humility, by working hard to educate ourselves and others about the innovation, understand its effects on consumers and the marketplace, identify benefits and likely harms, and, if harms do arise, consider whether existing laws and regulations are sufficient to address them, before assuming that new rules are required.[210]

This position does not mean there is no role for government as it pertains to artificial intelligence technologies. But it does mean that policymakers should first seek less restrictive remedies to complex social problems before resorting to policy proposals that are preemptive, proscriptive, and top-down in nature.

Permissionless Innovation as the Default Policy Position

What, then, should policymakers do? First, regulators and policymakers must carefully ensure that they have a full understanding of the boundaries and promises of all the technologies and applications that they address. A recent Mercatus Center report offered up a 10-part checklist that policymakers could use to help spur the development of dynamic new sectors and technologies.[211] That blueprint is as follows:

1. Articulate and defend permissionless innovation as the general policy default.

210. Maureen K. Ohlhausen, "The Internet of Things and the FTC: Does Innovation Require Intervention?" (remarks before the US Chamber of Commerce, Washington, DC, October 18, 2013).
211. Adam Thierer and Michael Wilt, "Permissionless Innovation: A 10-Point Checklist for Public Policymakers," Mercatus Center at George Mason University, March 31, 2016.

2. Identify and remove barriers to entry and innovation.
3. Protect freedom of speech and expression.
4. Retain and expand immunities for intermediaries from liability associated with third-party uses.
5. Rely on existing legal solutions and the common law to solve problems.
6. Wait for insurance markets and competitive responses to develop.
7. Push for industry self-regulation and best practices.
8. Promote education and empowerment solutions and be patient as social norms evolve to solve challenges.
9. Adopt targeted, limited legal measures for truly hard problems.
10. Evaluate and reevaluate policy decisions to ensure they pass a strict benefit-cost analysis.

To the extent that policymakers wish to spur the development of a wide array of new life-enriching technologies while also looking to devise sensible solutions to complex challenges, policymakers should consider this sort of flexible, bottom-up approach as the basis of America's policy regime for artificial intelligence systems and technologies.

Many AI technologies pose little or no risks to safety, fair market competition, or consumer welfare. These applications should not be stymied by an inappropriate and ill-defined regulatory scheme that seeks to address an entirely separate technology. They should be distinguished and exempted from regulations as appropriate.

Other AI technologies may warrant more regulatory consideration if they generate substantial risks to public welfare.[212] But preemptive policy prohibitions on innovation should be viewed as a last resort. Instead, the same sort of multistakeholder "soft law" approach that has guided internet policy over the past two decades can also serve as a governance model for AI. Indeed, this approach is perhaps the likeliest way forward for much AI oversight.[213]

212. This may be the case when the threat of harm associated with some technology is found to be highly probable, tangible, immediate, irreversible, and catastrophic. For more information, see Adam Thierer, "Wendell Wallach on the Challenge of Engineering Better Technology Ethics," *Technology Liberation Front*, April 20, 2016.
213. Kate Crawford et al., "The AI Now Report: The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term" (summary report of the AI Now public symposium hosted by the White House and New York University's Information Law Institute, July 7, 2016, New York, September 22, 2016), 24–25.

Soft Law Alternatives

Perhaps the most important of these recommendations is the call to rely on self-regulation, best practices, and other codes of conduct or developer guidelines to address AI-related concerns. And that appears to be the direction some policymakers are heading already.

Rather than the outdated command-and-control model for industry regulation that dominated in the wake of the Great Depression, regulators now more often turn to a hybrid type of oversight sometimes called "multistakeholderism." This approach to regulation eschews top-down precautionary controls (or "hard governance") in favor of a collaborative and informal governance structure undertaken by a variety of government, nonprofit, and industry bodies (also known as "soft governance").[214] Soft law approaches will likely be increasingly essential because, as University of Ottawa ethical philosopher Marc A. Saner points out, "the control paradigm is too limited to address all the issues that arise in the context of emerging technologies."[215] By the control paradigm, he generally means traditional administrative regulatory agencies and processes. He and other contributors to a 2014 book all seem to agree that the control problem paradigm "has its limits when diffusion, pacing and ethical issues associated with emerging technologies become significant, as is often the case."[216]

How, specifically, would such a soft governance system administer policy? Not with mandates and directives but with negotiated codes of conduct, voluntary best practices, and industry guidance and consultation.[217] Experts in technology and policy from a variety of institutions—academic, commercial, and governmental—are brought together to monitor threats and proactively consider solutions. Voluntary collaboration among these disparate groups, not hierarchical threats of control or prohibition, drives action.

In fact, many soft governance activities are in quiet but effective operation today. The Department of Commerce (the National Telecommunications

214. See Adam Thierer, "Innovation Arbitrage, Technological Civil Disobedience, and Spontaneous Deregulation," *Technology Liberation Front*, December 5, 2016.
215. Marc A. Saner, "The Role of Adaptation in the Governance of Emerging Technologies," in *Innovative Governance Models for Emerging Technologies*, ed. Gary E. Marchant, Kenneth W. Abbott, and Braden Allenby (Cheltenham, UK: Edward Elgar, 2014), 106.
216. Ibid.
217. As Wendell Wallach and Gary E. Marchant describe, these mechanisms include "codes of conduct, statements of principles, partnership programs, voluntary programs and standards, certification programs and private industry initiatives." See Gary E. Marchant and Wendell Wallach, "Governing the Governance of Emerging Technologies," in Marchant, Abbott, and Allenby, *Innovative Governance Models for Emerging Technologies*.

and Information Administration, in particular) and the FTC have already developed many industry codes of conduct and best practices for technologies such as biometrics,[218] big data,[219] the internet of things,[220] online advertising,[221] and much more.

A recent report from a public symposium of AI experts hosted by the White House and New York University's Information Law Institute noted that many existing professional organizations have professional codes of ethics that might be applicable to many policy concerns.[222] The organizations include the Association for the Advancement of Artificial Intelligence (AAAI), the Association of Computing Machinery (ACM), and the Institute of Electrical and Electronics Engineers (IEEE). Additionally, a new industry group called the Partnership on AI to Benefit People and Society was founded in 2016 by major commercial players such as Apple, Microsoft, Amazon, Google, IBM, and Facebook, as well as the American Civil Liberties Union, to "discuss and provide guidance on emerging issues related to the impact of AI on society."[223] These and other organizations can update existing best practices in collaboration with government and industry groups.

The role of the International Organization for Standardization (ISO), a global standards-making body that was formed in 1946, is particularly important in this regard. The ISO "is an independent, non-governmental international organization with a membership of 163 national

"Many existing professional organizations have professional codes of ethics that might be applicable to many policy concerns."

218. Federal Trade Commission, *Facing Facts: Best Practices for Common Uses of Facial Recognition Technologies* (FTC staff report, October 22, 2012).
219. Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion?*
220. Federal Trade Commission, *Internet of Things: Privacy and Security in a Connected World* (FTC staff report, January 27, 2015).
221. Federal Trade Commission, *.com Disclosures: How to Make Effective Disclosures in Digital Advertising* (FTC staff guidance, March 2013).
222. Crawford et al., "AI Now Report," 24–25.
223. Partnership on AI to Benefit People and Society Mission Statement, accessed April 4, 2017, https://www.partnershiponai.org/#s-mission.

standards bodies"[224] that seeks to build global consensus through multistakeholder efforts.[225] This is accomplished through the work of dozens of technical committees made up of experts from across the globe in diverse fields: industry, consumer associations, academia, nongovernmental organizations, and governments.[226] Such efforts can assist in the development of globally recognized best practices for robotics and AI-based systems. For example, in 2014, the ISO issued a set of safety requirements for personal care robots that "specifies requirements and guidelines for the inherently safe design, protective measures, and information for use of personal care robots."[227] This set of requirements is one of two dozen robotics-related standards that the organization had published or was in the process of formulating as of the beginning of 2017.[228] The ISO also has previously published numerous standards of automation systems and integration that could have relevance here.

Other guidelines and codes of conduct have been developed or proposed for robotics and AI. In late 2016, the British Standards Institute published a "Guide to the Ethical Design and Application of Robots and Robotic Systems."[229] The guide, which was written by a committee of scientists, academics, ethicists, and philosophers, "recognizes that potential ethical hazards arise from the growing number of robots and autonomous systems being used in everyday life" and, therefore, "provides additional guidelines to eliminate or reduce the risks associated with these ethical hazards to an acceptable level. These cover safe design, protective measures and information for the design and application of robots" in use in fields that range from industrial to personal care to medical.[230]

Meanwhile in the United States, in its solicitation announcement for public comment, the White House announced the formation of a new National Science and Technology Council Subcommittee on Machine Learning and Artificial

---

224. "About ISO," International Organization for Standardization (ISO), accessed January 13, 2017, http://www.iso.org/iso/home/about.htm.
225. "Technical Committees," ISO, http://www.iso.org/iso/home/standards_development/list_of_iso_technical_committees.htm.
226. "How We Develop Standards?," ISO, accessed January 13, 2017, http://www.iso.org/iso/home/standards_development.htm.
227. "13482:2014: Robots and Robotic Devices—Safety Requirements for Personal Care Robots," ISO, February 2014, http://www.iso.org/iso/catalogue_detail?csnumber=53820.
228. "Standards Catalogue: ISO/TC 299—Robotics," ISO, accessed January 13, 2017, http://www.iso.org/iso/home/store/catalogue_tc/catalogue_tc_browse.htm?commid=5915511.
229. Hannah Devlin, "Do No Harm, Don't Discriminate: Official Guidance Issued on Robot Ethics," *Guardian*, September 18, 2016.
230. British Standards Institution, "Robots and Robotic Devices: Guide to the Ethical Design and Application of Robots and Robotic Systems" (guidelines for BS 8611:2016, April 2016).

Intelligence.[231] This body will reportedly "monitor state-of-the-art advances and technology milestones in artificial intelligence and machine learning within the Federal Government, in the private sector, and internationally; and help coordinate Federal activity in this space." With input from industry and academic researchers, regulators, and bodies such as the new subcommittee, this group should develop a proper system of classification that defines each of the many disparate technologies that fall under the artificial intelligence umbrella.

How, then, can policymakers best oversee and guide the efforts of these disparate soft governance groups? Gary E. Marchant and Wendell Wallach propose the formation of what they call governance coordinating committees (GCCs) to work together with all the interested stakeholders to monitor technological development and to develop solutions to perceived problems. Rather than overlapping with or functioning as a regulatory body, the committee would work together with existing institutions.[232]

If done properly, GCCs, or something like them, could provide appropriate counsel and recommendations without the often-onerous costs of traditional regulatory structures. Additionally, private-public education and empowerment-based strategies could help the public learn to cope with new innovations or use them appropriately. And there are many flexible, ex post remedies to assist if things should go wrong.[233] For example, common law remedies such as product defects law, torts, contract law, property law, and even class action lawsuits have augmented traditional regulation in the past.[234]

Soft law approaches may lead to potential problems, however. Without appropriate and clearly defined boundaries and oversight, a well-intentioned soft governance system could ultimately evolve into a kind of soft tyranny that uses strong-armed agency threats to accomplish policy priorities outside the boundaries of clear legislative authority.[235] For example, such bodies could develop into a kind of favor-trading network, in which the interests of the strongest dominate. Of course, these are the same threats facing traditional regulatory structures.[236]

---

231. Ed Felten, "Preparing for the Future of Artificial Intelligence."
232. Marchant and Wallach, "Governing the Governance of Emerging Technologies."
233. For one early theoretical construct showing how tort law can remediate harms from AI technology, see William D. Smart, Cindy M. Grimm, and Woodrow Hartzog, "An Education Theory of Fault for Autonomous Systems" (white paper, presented to We Robot 2017, Information Society Project, Yale Law School, March 31–April 1, 2017).
234. Thierer, *Permissionless Innovation*, 120–25.
235. Jerry Brito, "'Agency Threats' and the Rule of Law: An Offer You Can't Refuse," *Harvard Journal of Law and Public Policy* 37, no. 2 (2014).
236. George J. Stigler, "The Theory of Economic Regulation," *Bell Journal of Economics and Management Science* 2, no. 1 (1971).

On the one hand, GCC-like bodies would not be elevated to the same level as an administrative regulating body, so their formal power would be limited. On the other hand, the informal nature of such bodies might lend themselves to less-transparent, and potentially harder-to-reverse, methods of capture.

CONCLUSION

Stanford University recently brought together 17 of the leading experts on AI issues to compile a comprehensive report, *The One Hundred Year Study on Artificial Intelligence*, which was billed as "a long-term investigation of the field of Artificial Intelligence (AI) and its influences on people, their communities, and society." The group's final report, which was published in September 2016, concluded by noting that

> Misunderstanding about what AI is and is not, especially against a background of scare-mongering, could fuel opposition to technologies that could benefit everyone. This would be a tragic mistake. Regulation that stifles innovation, or relocates it to other jurisdictions, would be similarly counterproductive.[237]

We concur. Toward that end, policymakers considering the best ways to oversee nascent AI technologies should build on the success with soft law and multistakeholder-style informal oversight mechanisms and eschew the prohibitive and innovation-limiting precautionary regulation of the past. AI technologies offer many challenges to overcome but an incredible array of promising applications, economic opportunities, and improvements to quality of life should we get our policies right. The benefits of AI technologies are simply too great for us to allow them to be extinguished by poorly considered policy.

237. Stone et al., *Artificial Intelligence and Life in 2030*.

ABOUT THE AUTHORS

Adam Thierer is a senior research fellow with the Technology Policy Program at the Mercatus Center at George Mason University. He specializes in technology, media, internet, and free-speech policies. His writings have appeared in the *Wall Street Journal*, the *Economist*, and the *Washington Post*, and his latest book is *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom*. Thierer has served on several distinguished online safety task forces, including Harvard University's Internet Safety Technical Task Force and the federal government's Online Safety Technology Working Group. Previously, he was president of the Progress & Freedom Foundation, director of telecommunications studies at the Cato Institute, and a senior fellow at the Heritage Foundation. Thierer received his MA in international business management and trade theory at the University of Maryland.

Andrea Castillo O'Sullivan is the program manager of the Technology Policy Program at the Mercatus Center and is pursuing her PhD in economics at George Mason University. Her research focuses on cybersecurity, government surveillance, internet freedom, cryptocurrency, and the economics of technology. She is a coauthor of *Liberalism and Cronyism: Two Rival Political and Economic Systems* with Randall G. Holcombe and of *Bitcoin: A Primer for Policymakers* with Jerry Brito. O'Sullivan received her BS in economics and political science from Florida State University.

Raymond Russell was a 2016 Google Policy Fellow at the Mercatus Center. His research interests include data science and the economics of technological change. He is an undergraduate at the University of Washington studying physics and economics.

ABOUT THE MERCATUS CENTER AT GEORGE MASON UNIVERSITY

The Mercatus Center at George Mason University is the world's premier university source for market-oriented ideas—bridging the gap between academic ideas and real-world problems.

A university-based research center, Mercatus advances knowledge about how markets work to improve people's lives by training graduate students, conducting research, and applying economics to offer solutions to society's most pressing problems.

Our mission is to generate knowledge and understanding of the institutions that affect the freedom to prosper and to find sustainable solutions that overcome the barriers preventing individuals from living free, prosperous, and peaceful lives.

Founded in 1980, the Mercatus Center is located on George Mason University's Arlington and Fairfax campuses.