Congressional Research Service
Informing the legislative debate since 1914

Artificial Intelligence: Background, Selected Issues, and Policy Considerations

May 19, 2021

Congressional Research Service
https://crsreports.congress.gov
R46795

SUMMARY

R46795

May 19, 2021

Laurie A. Harris
Analyst in Science and Technology Policy

Artificial Intelligence: Background, Selected Issues, and Policy Considerations

The field of artificial intelligence (AI)—a term first used in the 1950s—has gone through multiple waves of advancement over the subsequent decades. Today, AI can broadly be thought of as computerized systems that work and react in ways commonly thought to require intelligence, such as the ability to learn, solve problems, and achieve goals under uncertain and varying conditions. The field encompasses a range of methodologies and application areas, including machine learning (ML), natural language processing, and robotics.

In the past decade or so, increased computing power, the accumulation of big data, and advances in AI techniques have led to rapid growth in AI research and applications. Given these developments and the increasing application of AI technologies across economic sectors, stakeholders from academia, industry, and civil society have called for the federal government to become more knowledgeable about AI technologies and more proactive in considering public policies around their use.

Federal activity addressing AI accelerated during the 115th and 116th Congresses. President Donald Trump issued two executive orders, establishing the American AI Initiative (E.O. 13859) and promoting the use of trustworthy AI in the federal government (E.O. 13960). Federal committees, working groups, and other entities have been formed to coordinate agency activities, help set priorities, and produce national strategic plans and reports, including an updated *National AI Research and Development Strategic Plan* and a *Plan for Federal Engagement in Developing Technical Standards and Related Tools in AI*. In Congress, committees held numerous hearings, and Members introduced a wide variety of legislation to address federal AI investments and their coordination; AI-related issues such as algorithmic bias and workforce impacts; and AI technologies such as facial recognition and deepfakes. At least four laws enacted in the 116th Congress focused on AI or included AI-focused provisions.

- The National Defense Authorization Act for FY2021 (P.L. 116-283) included provisions addressing various defense- and security-related AI activities, as well as the expansive National Artificial Intelligence Initiative Act of 2020 (Division E).
- The Consolidated Appropriations Act, 2021 (P.L. 116-260) included the AI in Government Act of 2020 (Division U, Title I), which directed the General Services Administration to create an AI Center of Excellence to facilitate the adoption of AI technologies in the federal government.
- The Identifying Outputs of Generative Adversarial Networks (IOGAN) Act (P.L. 116-258) supported research on Generative Adversarial Networks (GANs), the primary technology used to create deepfakes.
- P.L. 116-94 established a financial program related to exports in AI among other areas.

AI holds potential benefits and opportunities, but also challenges and pitfalls. For example, AI technologies can accelerate and provide insights into data processing; augment human decisionmaking; optimize performance for complex tasks and systems; and improve safety for people in dangerous occupations. On the other hand, AI systems may perpetuate or amplify bias, may not yet be fully able to explain their decisionmaking, and often depend on vast datasets that are not widely accessible to facilitate research and development (R&D). Further, stakeholders have questioned the adequacy of human capital in both the public and private sectors to develop and work with AI, as well as the adequacy of current laws and regulations for dealing with societal and ethical issues that may arise. Together, such challenges can lead to an inability to fully assess and understand the operations and outputs of AI systems—sometimes referred to as the “black box” problem.

Because of these questions and concerns, some stakeholders have advocated for slowing the pace of AI development and use until more research, policymaking, and preparation can occur. Others have countered that AI will make lives safer, healthier, and more productive, so the federal government should not attempt to slow it, but rather should give broad support to AI technologies and increase federal AI funding.

In response to this debate, Congress has begun discussing issues such as the trustworthiness, potential bias, and ethical uses of AI; possible disruptive impacts to the U.S. workforce; the adequacy of U.S. expertise and training in AI; domestic and international efforts to set technological standards and testing benchmarks; and the level of U.S. federal investments in AI research and development and how that impacts U.S. international competitiveness. Congress is likely to continue grappling

with these issues and working to craft policies that protect American citizens while maximizing U.S. innovation and leadership.

Contents

Figures

Contacts

Introduction

Artificial intelligence (AI)—a term first used in the 1950s—can broadly be thought of as computerized systems that work and react in ways commonly thought to require intelligence, such as the ability to learn, solve problems, and achieve goals under uncertain and varying conditions.[1] In the past decade, increases in computing power, the availability of large-scale datasets (i.e., big data), and advances in the methodologies underlying AI, have led to rapid growth in the field. AI technologies currently show promise for improving the safety, quality, and efficiency of work and for promoting innovation and economic growth. At the same time, the application of AI to complex problem solving in real-world situations raises concerns about trustworthiness, bias, and ethics and potential disruptive effects on the U.S. workforce. In addition, numerous policy questions are at issue, including those concerning the appropriate U.S. approach to international competition in AI research and development (R&D), technological standard setting, and the development of testing benchmarks.

Given the increasing use of AI technologies across economic sectors, stakeholders from academia, industry, and civil society have called for the federal government to become more knowledgeable about AI technologies and more proactive in considering public policies around their use. To assist Congress in its work on AI, this report provides an overview of AI technologies and their development, recent trends in AI, federal AI activity, and selected issues and policy considerations.

This report does not attempt to address all applications of AI. Information on the application of AI technologies in transportation, national security, and education can be found in separate CRS products.[2]

What Is AI?

While there is no single, commonly agreed upon definition of AI, the National Institute of Standards and Technology (NIST) has described AI technologies and systems as comprising "software and/or hardware that can learn to solve complex problems, make predictions or undertake tasks that require human-like sensing (such as vision, speech, and touch), perception, cognition, planning, learning, communication, or physical action."[3] Definitions may vary according to the discipline in which AI is being discussed.[4] AI is often described as a field that encompasses a range of methodologies and application areas, such as machine learning (ML), natural language processing (NLP), and robotics.

[1] Adapted from Office of Science and Technology Policy, *Preparing for the Future of Artificial Intelligence*, October 2016, p. 6.

[2] See CRS Report R44940, *Issues in Autonomous Vehicle Deployment*, by Bill Canis; CRS In Focus IF10737, *Autonomous and Semi-autonomous Trucks*, by John Frittelli; CRS Report R45178, *Artificial Intelligence and National Security*, by Kelley M. Sayler; and CRS In Focus IF10937, *Artificial Intelligence (AI) and Education*, by Joyce J. Lu and Laurie A. Harris.

[3] National Institute of Standards and Technology, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, August 9, 2019, pp. 7-8.

[4] See, for example, AI definitions in the categories of ordinary language, computational disciplines, engineering, economics and social sciences, ethics and philosophy, and international law and policy, in Sara Mattingly-Jordan et al., *Ethically Aligned Design: First Edition Glossary*, Institute of Electrical and Electronics Engineers (IEEE), January 2019, p. 8, at https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_glossary.pdf.

Defining AI is not merely an academic exercise, particularly when drafting legislation. AI research and applications are evolving rapidly. Thus, congressional consideration of whether to include a definition for AI in a bill, and if so how to define the term or related terms, necessarily include attention to the scope of the legislation and the current and future applicability of the definition. Considerations in crafting a definition for use in legislation include whether it is expansive enough not to hinder the future applicability of a law as AI develops and evolves, while being narrow enough to provide clarity on the entities the law affects. Some stakeholders, recognizing the many challenges of defining AI, have attempted to define principles that might help guide policymakers. Research suggests that differences in definitions used to identify AI-related research may contribute to significantly different analyses and outcomes regarding AI competition, investments, technology transfer, and application forecasts.[5]

The John S. McCain National Defense Authorization Act for Fiscal Year 2019 (P.L. 115-232) included the first definition of AI in federal statute.[6] Like those in other previously introduced bills, the definition incorporated a commonly cited framework of four possible goals that AI systems may pursue: systems that think like humans (e.g., neural networks), act like humans (e.g., natural language processing), think rationally (e.g., logic solvers), or act rationally (e.g., intelligent software agents embodied in robots).[7] However, AI research and applications do not necessarily fall solely within any one of these four categories.

In December 2020, the National Artificial Intelligence Act of 2020, enacted as part of the William M. (Mac) Thornberry National Defense Authorization Act (NDAA) for Fiscal Year 2021 (P.L. 116-283), included the following definition:

> The term "artificial intelligence" means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to—(A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action.[8]

Current AI systems are considered to be *narrow AI*, meaning that they are tailored to particular, narrowly defined tasks. Example applications of AI in everyday life include email spam filtering, voice assistance (e.g., Siri, Alexa, Cortana), financial lending decisions, and search engine results. AI technologies are being integrated in a range of sectors, including transportation, health care, education, agriculture, manufacturing, and defense. Some AI experts use the terms *augmented intelligence* or *human-centered AI* to capture the various AI applications in physical and connected systems, such as robotics and the Internet of Things,[9] and to emphasize the use of AI technologies to enhance human activities rather than to replace them.

Most analysts believe that *general AI*, meaning systems that demonstrate intelligent behavior across a range of cognitive tasks, is unlikely to occur for a decade or longer. Some AI researchers

[5] Dewey Murdick, James Dunham, and Jennifer Melot, *AI Definitions Affect Policymaking*, Center for Security and Emerging Technology, June 2020, at https://cset.georgetown.edu/wp-content/uploads/CSET-AI-Definitions-Affect-Policymaking.pdf.

[6] P.L. 115-232, Section 238; 10 U.S.C. §2358 note.

[7] Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. (Upper Saddle River, NJ: Prentice Hall, 2010), pp. 1-5.

[8] P.L. 116-283 (hereinafter referred to as the FY2021 NDAA); H.R. 6395, Division E, Section 5002(3).

[9] For more information on the Internet of Things, see CRS In Focus IF11239, *The Internet of Things (IoT): An Overview*, by Patricia Moloney Figliola; and to identify additional CRS experts who work on IoT and related topics, see CRS Report R44225, *The Internet of Things: CRS Experts*, coordinated by Patricia Moloney Figliola.

believe that general AI can be achieved through incremental development and refining of current AI and machine learning tools, while others believe it will require the discovery and development of a new breakthrough technique.

Just as there is debate over the definition of AI, there is debate over which technologies should be classified as AI. For example, robotic process automation (RPA) has been defined as "the use of software to automate highly repetitive, routine tasks normally performed by knowledge workers."[10] Because it automates activities performed by humans, it is often described as an AI technology. However, some argue that RPA is not AI because it does not include a learning component. Others discuss RPA as a basic tool that can be combined with AI to create complex process automation (CPA) or intelligent process automation (IPA), along an "intelligent automation continuum."[11]

AI Terminology

Some stakeholders, including industry, advocacy groups, and policymakers, have raised questions about whether specific AI technologies and techniques require tailored legislation. For example, legislation enacted in the 116th Congress focused on generative adversarial networks (GANs), described below, which are the main underlying AI technique used in generating deepfakes,[12] which are most commonly described as realistic audio, video, and other forgeries created using AI techniques.[13] This section is meant to provide a broad understanding of a subset of common terms used in the field of AI and how they relate to one another. These include the subfield of machine learning (ML); ML techniques such as deep learning, neural networks, and GANs; and training methods such as supervised, unsupervised, and reinforcement learning. However, just as there are variations in how AI is defined, researchers and practitioners describe various AI-related terms in slightly different ways. Further, the following terms and techniques are not mutually exclusive; AI systems may employ more than one. For example, AlphaGo—the first AI program to beat a human master at the ancient Chinese game of Go—combined deep neural networks, supervised learning, and reinforcement learning.[14]

- **Machine learning** (ML), often referred to as a subfield of AI, examines how to build computer programs that automatically improve their performance at some task through experience without relying on explicit rules-based programming to do so.[15] One of the goals of ML is to teach algorithms to successfully interpret data that have not previously been encountered. ML is one of the most common AI techniques in use today, and most ML tasks are narrowly specified to optimize

[10] See IBM, "Automate Repetitive Tasks," at https://www.ibm.com/automation/rpa.

[11] IBM Global Business Services, "Using Artificial Intelligence to Optimize the Value of Robotic Process Automation," September 2017, at https://www.ibm.com/downloads/cas/KDKAAK29.

[12] The Identifying Outputs of Generative Adversarial Networks (IOGAN) Act (P.L. 116-258).

[13] For additional information on deepfakes, see CRS In Focus IF11333, *Deep Fakes and National Security*, by Kelley M. Sayler and Laurie A. Harris.

[14] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. (Cambridge, MA: MIT Press, 2018), pp. 441-442.

[15] Adapted from Erik Brynjolfsson, Tom Mitchell, and Daniel Rock, "What Can Machines Learn, and What Does It Mean for Occupations and the Economy?," *AEA Papers and Proceedings*, vol. 108 (May 1, 2018), pp. 43-47, at http://www-cgi.cs.cmu.edu/~tom/pubs/AEA2018-WhatCanMachinesLearn.pdf. ML is defined in P.L. 116-293 to mean "an application of artificial intelligence that is characterized by providing systems the ability to automatically learn and improve on the basis of data or experience, without being explicitly programmed."

specific functions using particular datasets. Deep learning, neural networks, and GANs represent a few of the ML techniques frequently used today.

- **Deep learning** (DL) systems learn from large amounts of data to subsequently recognize and classify related, but previously unobserved, data. For example, **neural networks**, often described as being loosely modeled after the human brain, consist of thousands or millions of processing nodes generally organized into layers. The strength of the connections among nodes and layers are repeatedly tuned—based on characteristics of the training data—to correspond to the correct output. Advances in hardware, such as the development of graphical processing units (GPUs), have allowed these networks to have many layers, which is what puts the "deep" in deep learning. DL approaches have been used in systems across many areas of AI research, from autonomous vehicles to voice recognition technologies.[16]
- **Generative adversarial networks** (GANs) consist of two competing neural networks—a generator network that tries to create fake outputs (such as pictures), and a discriminator network that tries to determine whether the outputs are real or fake. A major advantage of this structure is that GANs can learn from less data than other deep learning algorithms.[17] Adversarial ML systems can be used in other ways, as well; for example, they can try to improve fairness in financial service decisionmaking by having a second model try to guess the protected class of applicants based on models built by another model.[18]
- **Supervised learning** algorithms learn from a training set of data that is labeled with the correct description (e.g., the correct label for this picture is "cat"); the system subsequently learns which components of the data are useful for classifying it correctly and uses that information to correctly classify data it has never encountered before. In contrast, **unsupervised learning** algorithms search for underlying structures in unlabeled data.
- **Reinforcement learning** (RL) refers to giving computer programs the ability to learn from experience, providing them with minimal inputs and human interventions.[19] RL algorithms learn by trial and error, being rewarded for reaching specified objectives—both for immediate actions and long-term goals. The emphasis on simulated motivation and learning from direct interaction with the environment, without requiring explicit examples and models, are among the characteristics that set RL apart from other ML approaches.[20]

---

[16] Larry Hardesty, "Explained: Neural Networks," *Massachusetts Institute of Technology (MIT) News*, April 14, 2017, at http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414.

[17] Jamie Beckett, "What's a Generative Adversarial Network? Leading Researcher Explains," *NVIDIA*, May 17, 2017, at https://blogs.nvidia.com/blog/2017/05/17/generative-adversarial-network/.

[18] Sally Ward-Foxton, "Reducing Bias in AI Models for Credit and Loan Decision," *EE Times*, April 30, 2019, at https://www.eetimes.com/reducing-bias-in-ai-models-for-credit-and-loan-decisions/#.

[19] Sean Garrish, *How Smart Machines Think* (Cambridge, MA: MIT Press, 2018), p. 91.

[20] Adapted from Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. (Cambridge, MA: MIT Press, 2018).

Algorithms and AI

As interest in AI continues to grow, some analysts assert that general data analytics and specialized algorithms are increasingly being referred to, erroneously, as AI. It can be challenging to make such distinctions clearly, given the variability in definitions of AI and related terms and because these distinctions have arguably shifted over time. For example, an algorithm is basically a procedure or set of instructions designed to perform a specific task or solve a mathematical problem. Some early products of AI research, such as rule-based expert systems, are algorithms encoded with expert knowledge but lacking a learning component. Some feel that rule-based systems are a simple form of AI because they simulate intelligence, while others think that without a learning component a system should not be considered AI.[21] Generally, however, the goals of AI—automating or replicating intelligent behavior—have remained consistent.[22]

Historical Context of AI

The ideas underlying AI and its conceptual framework have been researched since at least the 1940s and initially formalized in the 1950s. Ideas about intelligent machines were discussed and popularized by scientists and authors such as Alan Turing and Isaac Asimov,[23] and the term "artificial intelligence" was coined at the Dartmouth Summer Research Project on Artificial Intelligence, proposed in 1955 and held the following year.[24]

Since that time, the field of AI has gone through what have been termed by some as summers and winters—periods of much research and advancement, followed by lulls in activity and progress. The reasons for the AI winters have included a focus on theory over practical applications, research problems being more difficult than anticipated, and limitations of the technologies of the time. Much of the current progress and research in AI, which began around 2010, has been attributed to the availability of big data, improved ML approaches and algorithms, and more powerful computers.[25]

Waves of AI

The Defense Advanced Research Projects Agency (DARPA), which has funded AI R&D since the 1960s, has described the development of AI technologies in terms of three waves.[26] These waves are described by the varying abilities of technologies in each to *perceive* rich, complex, and subtle

[21] For a brief discussion see, for example, Tricentis, "AI Approaches Compared: Rule-Based Testing vs. Learning," at https://www.tricentis.com/artificial-intelligence-software-testing/ai-approaches-rule-based-testing-vs-learning/.

[22] Office of Science and Technology Policy, *Preparing for the Future of Artificial Intelligence*, October 2016, pp. 5-6.

[23] Alan M. Turing, "Computing Machinery and Intelligence," *Mind*, vol. 49 (1950), pp. 433-460, at https://www.csee.umbc.edu/courses/471/papers/turing.pdf; and Isaac Asimov, *I, Robot* (Garden City, NY: Doubleday, 1950).

[24] See J. McCarthy et al., "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," August 31, 1955, at http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

[25] Executive Office of the President, National Science and Technology Council, Committee on Technology, *Preparing for the Future of Artificial Intelligence*, October 2016, pp. 5-6.; for additional information on these factors and a short history of AI, see also the appendix of 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 2016, at http://ai100.stanford.edu/2016-report.

[26] See "DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies," September 7, 2018, at https://www.darpa.mil/news-events/2018-09-07.

information; to *learn* within an environment; to *abstract* to create new meanings; and to *reason* in order to plan and reach decisions.[27]

First wave: handcrafted knowledge. The first wave of AI technologies have abilities primarily to perceive and reason but no learning capability and poor handling of uncertainty. For such technologies, researchers and engineers create sets of rules to represent knowledge in well-defined domains for narrowly defined problems. The TurboTax software, an expert system, is one example. Rules are built into the application, which then turns input information into tax form outputs, but it has only a rudimentary ability to perceive and no ability to learn (e.g., about a new tax law) or to abstract beyond what it is programmed to know.

Second wave: statistical learning. Starting in the 1990s, a second wave of AI technologies were developed with more nuanced abilities to perceive and learn, with some ability to abstract, minimal reasoning ability, but no contextual ability. For these systems, engineers create statistical models for specific problem domains and train them on big data. Generally, while such systems are statistically powerful, they can be individually unreliable, especially in the presence of skewed training data (e.g., a face recognition system trained on a limited range of skin tones can be powerful for similar faces, but highly unreliable for individuals outside of the training spectrum). As noted by DARPA, these technologies are "dependent on large amounts of high quality training data, do not adapt to changing conditions, offer limited performance guarantees, and are unable to provide users with explanations of their results."[28] Additional examples of second wave AI technologies include voice recognition and text analysis.

Third wave: contextual adaptation. The third wave of AI technologies is oriented toward making it possible for machines to adapt to changing situations (i.e., contextual adaptation). Engineers create systems that construct explanatory models of real world phenomena, and "AI systems learn and reason as they encounter new tasks and situations."[29] Examples of third wave technologies would include explainable AI (XAI), as described below.

Recent Growth in the Field of AI

There are many potential indicators of growth in the AI field. This section presents indicators of growth based on R&D activities and public and private investments in areas of frequent congressional interest. It also provides a brief discussion of AI hype versus the reality of what AI technologies are capable of today. It can be challenging to obtain comprehensive and directly comparable data for the indicators discussed in this section, particularly for AI investments. Therefore, such data should be evaluated carefully and treated as only indicative of trends.

AI Research and Development

One way to assess the growth in AI R&D is based on the publication of peer-reviewed papers, including both conference papers and journal articles. According to the AI Index group, between 2000 and 2019, the total number of peer-reviewed AI publications in Elsevier's Scopus

[27] Arati Prabhakar, former Director of DARPA, "Powerful but Limited: A DARPA Perspective on AI," presentation at National Academies of Sciences, Engineering, and Medicine workshop, *Robotics and Artificial Intelligence: Policy Implications for the Next Decade*, December 12, 2016, at https://www.nationalacademies.org/event/12-12-2016/robotics-and-artificial-intelligence-policy-implications-for-the-next-decade (hereinafter "Prabhakar, 2016").

[28] See "DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies," September 7, 2108, at https://www.darpa.mil/news-events/2018-09-07.

[29] Prabhakar, 2016.

database—the world's largest abstract and citation database—grew nearly 12-fold.[30] Authors based in the Europe Union (EU) published the most peer-reviewed AI publications as a percentage of the world total from 2000 to 2007 and again from 2012 to 2016, while authors based in China published the most from 2008 to 2011 and 2017 to 2019.[31] In 2020, the papers published by authors in China surpassed those of authors in the United States in the share of AI journal citations in the world for the first time. However, over the past decade, authors in the United States have consistently had more cited AI conference papers than authors based in China.[32] Further, the number of publications a researcher or country produces does not necessarily equate to scientific impact or research quality. As one researcher at the University of Oxford, UK, reportedly stated, "Just pumping out raw numbers of papers that don't have a lasting impact isn't really useful. It's more important to keep up with the technology frontier."[33] Such evaluations, however, do not discuss the finer points of which studies included teams of researchers from more than one country, raising the question of how to neatly attribute papers to regions, organizations, or funding sources.

In addition to published papers, many AI researchers in recent years have published preprint papers (submitted before peer review) to an online repository called arXiv (pronounced "archive"). As reported by the AI Index group, between 2015 and 2020, the total number of AI papers on arXiv increased over six-fold, with more growth in certain subcategories, providing a rough indication of areas of research activity across a range of AI subfields.[34] As of 2020, the most common subcategories of preprint papers were ML and computer vision (**Figure 1**).[35]

Groups like the AI Index have also attempted to measure progress in AI and its fields of study, though critics have categorized such efforts as reporting "trends in data that are related to AI" rather than tracking progress.[36] Further, recent research has raised concerns about the accuracy of reported improvements. By some measures, such as training time and cost, areas such as image classification have improved substantially.[37] By other measures, researchers assert that progress has come from tweaks, rather than core innovations, and some purported progress might not have taken place. For example, some researchers using meta-analyses of algorithms in various fields and applications—such as pruning algorithms used to make neural networks more efficient and information retrieval programs used in search engines—have found no clear evidence of performance improvements over the 10-year period from 2010 to 2019.[38]

[30] AI Index Steering Committee, *The AI Index 2021 Annual Report*, Human-Centered AI Institute, Stanford University, Stanford, CA, March 2021, p. 18 (hereinafter, "AI Index 2021"). The AI Index 2021 report authors provided this information only for the United States, China, and the Europe Union (EU), not individual countries within the EU.

[31] Ibid, p. 20.

[32] Ibid., p. 17.

[33] Neil Savage, "The Race to the Top Among the World's Leaders in Artificial Intelligence," *Nature Index*, December 9, 2020, at https://www.nature.com/articles/d41586-020-03409-8.

[34] Ibid., p. 32.

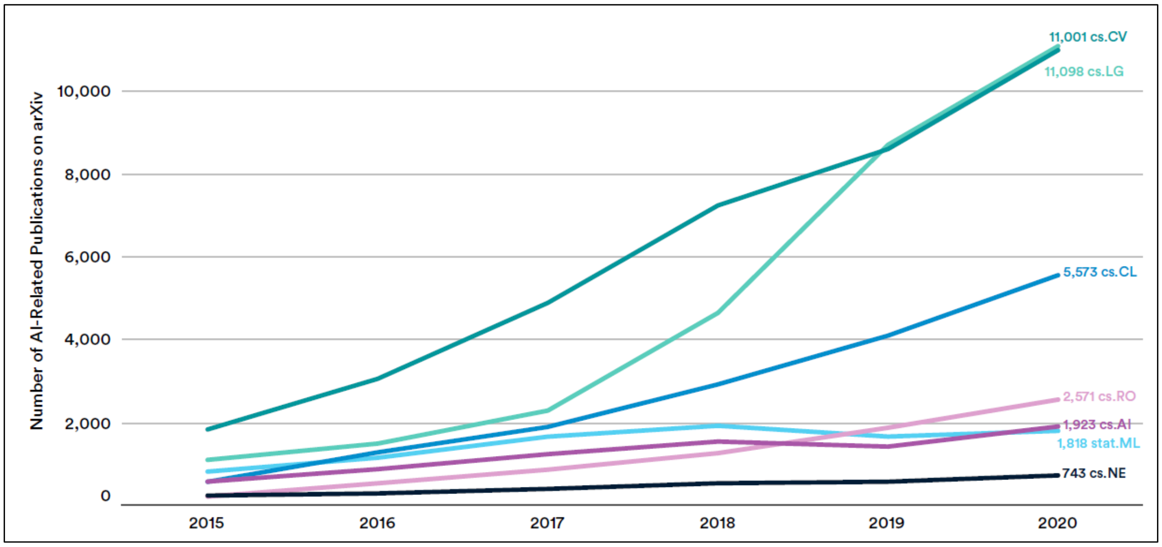
[35] Ibid., p. 34.

[36] Jeffrey Funk and Gary Smith, "Stanford's AI Index Report: How Much Is BS?," *Mind Matters News*, March 3, 2020, at https://mindmatters.ai/2020/03/stanfords-ai-index-report-how-much-is-bs/.

[37] AI Index 2021, pp. 48-49. Image classification broadly refers to the assigning of identification labels to images.

[38] Matthew Hutson, "Core Progress in AI Has Stalled in Some Fields," *Science*, vol. 368, no. 6494 (May 29, 2020), p. 927, at https://science.sciencemag.org/content/368/6494/927.

Figure 1. Total Number of AI-Related Publications on arXiv, by Field of Study, 2015-2020



Source: AI Index Steering Committee, *The AI Index 2021 Annual Report*, Human-Centered AI Institute, Stanford University, Stanford, CA, December 2021, p. 34.

Notes: The arXiv is an online repository for pre-publication papers, which generally means they have not undergone prior peer review. The papers on arXiv listed here are grouped by field of study, including cs.CV (computer vision), cs.LG (machine learning in computer science), cs.CL (computation and language), cs.RO (robotics), cs.AI (artificial intelligence), stat.ML (machine learning in statistics), and cs.NE (neural and evolutionary computing).

Private and Public Funding

Since around 2015, private funding for AI has been increasing, both in the United States and globally. For example, according to the AI Index 2021 report, global corporate investment in AI—including private investment, public offerings, mergers and acquisitions, and minority stakes—increased from $12.8 billion raised in 2015 to over $67.8 billion in 2020.[39] Global AI startup funding also increased steadily from 2015 to 2020, though the number of companies funded has decreased for each year from 2017 through 2020.[40] The United States continues to lead the world in private AI investments, with $23.6 billion in funding in 2020, followed by China ($9.9 billion) and the European Union ($2.0 billion). The top area of private investment in AI in 2020 was "Drugs, Cancer, Molecular, Drug Discovery" with more than $13.8 billion, 4.5 times higher than in 2019.[41] This increased funding for this particular area in 2020 may have been in large part a response to the Coronavirus Disease 2019 (COVID-19) pandemic; among the additional areas that also saw substantial increases in funding from 2019 to 2020 were "Students, Courses, Edtech, English language" and "Speech Recognition, Computer interaction, Dialogue, and Machine translation."[42] According to a McKinsey 2020 survey of over 1,000 company

[39] AI Index 2021, p. 93.

[40] Ibid, p. 94.

[41] Ibid, p. 11.

[42] Ibid, p. 97.

respondents, over half reported no change in AI investments amid the coronavirus pandemic, and 25% increased their investment in AI.[43]

In FY2020, U.S. public funding for AI R&D was reported for the first time across non-defense federal agencies in a supplemental report to the President's FY2020 budget, submitted by the Networking and Information Technology Research and Development (NITRD) Program. The annual NITRD supplemental report includes funding information across Program Component Areas (PCAs), which are major subject areas of federal IT R&D and may change each year. For FY2021, AI is included as a stand-alone PCA, and the report includes FY2019 actual investments, FY2020 enacted investments, and FY2021 requested funding amounts. While AI is a stand-alone PCA, some other PCAs have AI as a component.[44] Total FY2021 requested funding for non-defense agency AI R&D under the AI PCA was $912 million (an increase from the FY2020 enacted and supplemental total amount of $660 million); for AI-related efforts reported in other PCAs, the request was $590 million (an increase from the FY2020 enacted and supplemental total amount of $466 million). Thus, the total requested federal FY2021 non-defense budget for AI across PCAs was $1.5 billion (an increase from the FY2020 enacted and supplemental total amount of $1.1 billion).[45] By agency, the largest proportions of the FY2021 non-defense AI PCA request were from the National Science Foundation (NSF, $457 million), the U.S. Department of Agriculture (USDA, $128 million), and the Department of Energy (DOE, $84 million).[46]

Although defense agencies did not report AI funding numbers as part of the NITRD supplemental report, Bloomberg Government reported that the Department of Defense (DOD) FY2020 enacted budget for AI R&D was $5.0 billion, equal to the estimated FY2021 request.[47] The FY2021 request estimate included $568 million at DARPA, $250 million for the Algorithmic Cross Functional Team (also known as "Project Maven"), and $132 million for the Joint Artificial Intelligence Center (JAIC).[48]

Another measure of public investment in AI comes from data on government spending on AI contracts. According to analysis by Bloomberg Government using the Federal Procurement Data System (FPDS), in FY2018, U.S. federal agencies spent a total of $1.8 billion on unclassified AI-related contracts in FY2020, more than six times higher than the approximately $300 million spent in FY2015.[49] DOD accounts for the vast majority of FY2020 AI-related contract spending

---

[43] Tara Balakrishnan et al., *The State of AI in 2020*, McKinsey & Company, November 17, 2020, at https://www.mckinsey.com/Business-Functions/McKinsey-Analytics/Our-Insights/Global-survey-The-state-of-AI-in-2020. The survey and interviews with executives were conducted from May to August, 2020, and included 1,151 respondents from organizations that had adopted AI in at least one function out of a total of 2,395 participants.

[44] Examples of activities under the AI PCA include R&D that is primarily ML, and R&D focused on cybersecurity challenges unique to AI, and on computing architectures or chips optimized for neural networks. Examples of AI activities captured under other PCAs include R&D on robots that employ machine vision, R&D on the broad problem of human-machine interaction, and general research on neuromorphic computing. Ibid., pp. 11-12.

[45] Ibid., p. 11.

[46] Ibid., pp. 8-9; DOE NNSA is listed separately from DOE.

[47] As reported in AI Index 2021, p. 168.

[48] Ibid. Project Maven was launched in April 2017 and charged with rapidly incorporating AI into existing DOD systems to demonstrate the technology's potential; Robert Work, Former Deputy Secretary of Defense, Memorandum, "Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven)," April 26, 2017, at https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf. The JAIC is tasked with coordinating the efforts of DOD to develop, mature, and transition AI technologies into operational use, per P.L. 115-232, Section 2, Division A, Title X, §1051. Details and analysis for the FY2022 request are not yet available.

[49] As reported in AI Index 2021, p. 169.

($1.4 billion); after DOD, the National Aeronautics and Space Administration (NASA), the Department of Homeland Security (DHS), and the Department of Health and Human Services (HHS) have accounted for the largest share of spending on AI contracts among federal agencies since 2010.[50] FPDS data may be helpful in identifying broad trends and producing rough estimates, but as other analysts have noted, these data may not be reliable and decisionmakers should understand its limitations and be cautious in using the data to develop policy or draw conclusions.[51]

Important considerations in evaluating any of these numbers, and especially in attempting to compare them to funding amounts reported by other countries, are the various potential discrepancies in the numbers by year, investment type, and reporting entity. The AI Index group has previously asserted that there is no consensus on standard labeling for AI related investment activities, no existing measurement and accounting standards for public investment or expenditures in AI, and no consistently available data comparing public investments across countries.[52]

AI hype and reality. The recent growth and advances in the field of AI have been impressive, and notable researchers have highlighted both the far-reaching potential benefits, as well as the constraints and potential pitfalls of current and future AI technologies. Sergey Brin, co-founder of Google, has called the period of advancements over the past decade or so a "new spring in artificial intelligence," stating that we are in a "technology renaissance" with monthly advances and "applications across nearly every segment of modern society," while also highlighting potential concerns that accompany these advances (e.g., effects on employment, fairness, transparency, and safety).[53] AI systems currently remain constrained to narrowly-defined tasks and can fail with small modifications to inputs. For example, deep learning systems that have excelled at recognizing facial images can be deceived by the introduction of simple image distortions, or "noise" in the data.[54] The introduction of imperceptible or seemingly irrelevant changes to inputs, such as images, text, or sound waves, by malevolent actors has raised concerns about unforeseen vulnerabilities of AI, particularly in applications in autonomous vehicles, medical technologies, and defense systems. One expert noted, "While some people are worried about 'superintelligent' A.I., the most dangerous aspect of A.I. systems is that we will trust them too much and give them too much autonomy while not being fully aware of their limitations."[55] Many researchers agree that continued progress in AI requires the development and refinement of new techniques, in addition to increased availability of data and improvements in computing capacity.[56]

---

[50] Ibid.

[51] For additional discussions of FDPS data and how the FDPS system operates, see CRS Report R44010, *Defense Acquisitions: How and Where DOD Spends Its Contracting Dollars*, by John F. Sargent Jr. and Christopher T. Mann.

[52] AI Index Steering Committee, *The AI Index 2019 Annual Report*, Human-Centered AI Institute, Stanford University, Stanford, CA, December, 2019, p. 98.

[53] Sergey Brin, "2017 Founders' Letter," at https://abc.xyz/investor/founders-letters/2017/index.html.

[54] Gaurav Goswami et al., "Unravelling Robustness of Deep Learning Based Face Recognition Against Adversarial Attacks," *Association for the Advancement of Artificial Intelligence*, February 22, 2018, at https://arxiv.org/abs/1803.00401.

[55] Melanie Mitchell, "Artificial Intelligence Hits the Barrier of Meaning," *New York Times*, November 5, 2018, at https://www.nytimes.com/2018/11/05/opinion/artificial-intelligence-machine-learning.html.

[56] Tom Simonite, "Your Instagram #Dogs and #Cats Are Training Facebook's AI," *Wired*, May 2, 2018, at https://www.wired.com/story/your-instagram-dogs-and-cats-are-training-facebooks-ai/.

Selected Research and Focus Areas

AI research currently spans a broad range of techniques and application areas. This section describes a selection of areas that have received attention in recent years and may be of particular interest to Congress, including an example of AI in healthcare; it is not meant to portray any area as more or less valuable than another to the overall progress of AI research. Some of these areas include explainable AI, data access and models that can learn from reduced amounts of data, and hardware to improve the speed of, and reduce the computing power required to run, AI algorithms.

Explainable AI

As mentioned above in the discussion of third wave AI technologies, explainable AI has been an active area of research in recent years. As described by experts at DARPA, XAI research aims to create AI applications that can explain their actions and decisions to human users to improve trust and collaboration between humans and AI systems (**Figure 2**). Such explanations could help people identify and correct errors that AI systems make when generalizing from training data. This is of particular concern in high-stakes applications, such as classifying disease in medical images and classifying combatants and civilians in military surveillance images.[57]

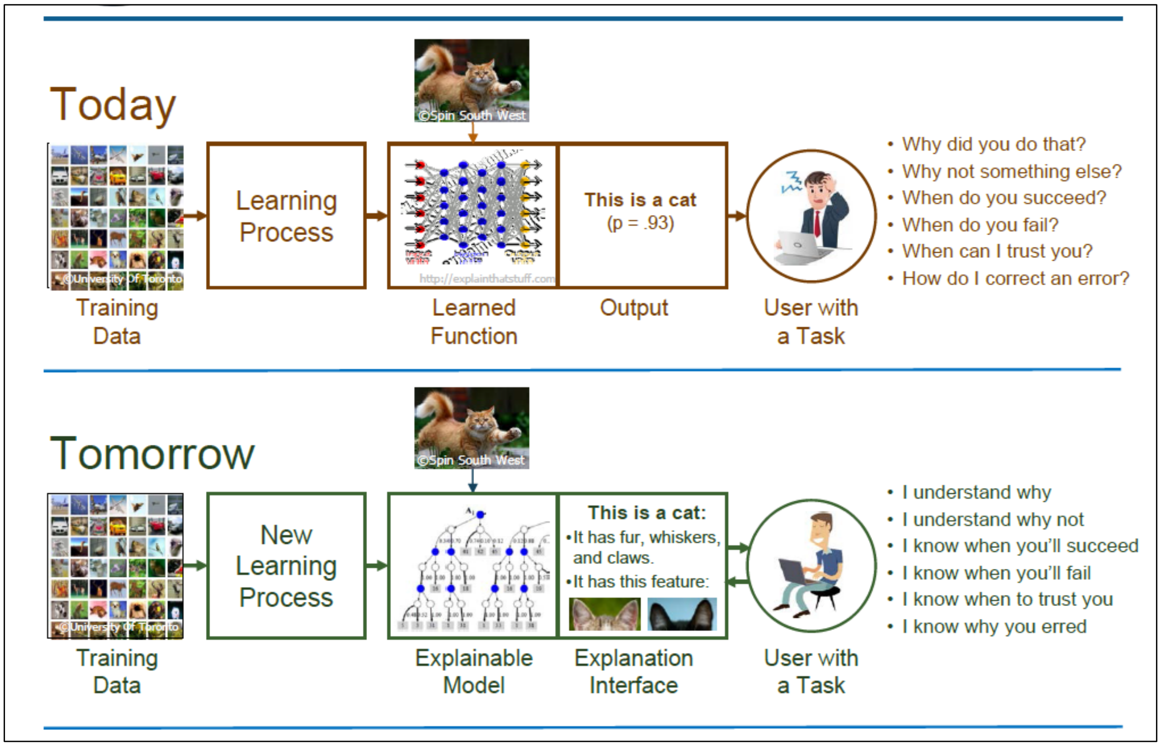
Federal agencies and the White House have been working to define and guide federal development and use of understandable and explainable AI systems. In August 2020, NIST released a draft publication for public comment on "Four Principles of Explainable Artificial Intelligence" that presents principles, categories, and theories of XAI.[58] In December 2020, Executive Order 13960 included, as a principle guiding the use of AI in federal government, that AI should be understandable, specifically that agencies shall "ensure that the operations and outcomes of their AI applications are sufficiently understandable by subject matter experts, users, and others."[59]

[57] For a deeper discussion of XAI, see also Alejandro Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI," *Information Fusion*, vol. 58 (June 2020), pp. 82-115, at https://www.sciencedirect.com/science/article/pii/S1566253519308103.

[58] National Institute of Standards and Technology, *Four Principles of Explainable Artificial Intelligence*, Draft NISTRIR 8312, August 2020, at https://www.nist.gov/document/four-principles-explainable-artificial-intelligence-nistir-8312.

[59] Executive Order 13960, "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government," 85 *Federal Register* 78939, December 3, 2020, at https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government.

Figure 2. Examples of Non-Explainable and Explainable AI Systems



Source: David Gunning, DARPA, "Explainable Artificial Intelligence (XAI) Program Update," November 2017, at https://web.archive.org/web/20200501004458/https://www.darpa.mil/attachments/XAIProgramUpdate.pdf.

Data Access

The availability of big data to train AI models enabled major advances in the field over the last decade. For example, the ImageNet project, which contains over 14 million publicly available labeled images, held competitions from 2010 through 2017 that led to improvements in AI visual recognition performance.[60] However, those developing AI technologies face barriers to using currently available datasets. In addition to the sheer amount of data available, researchers have noted the importance of using specific types of data of requisite quality for various applications of AI technologies, which can be expensive and time consuming to generate (e.g., data that have been digitally stored, cleaned, transformed, labeled, and optimized to be deployed in AI algorithms).[61] Associated data management infrastructure requirements can be extensive, including cloud technology, edge computing (computing done closer to the source of the data), and labeling and annotation capacity (human capital).[62]

While big data sets continue to be instrumental in various AI advances, some have raised concerns that such datasets are increasingly held by private companies and argued for more publicly available datasets and incentives for technology companies to share proprietary datasets. One study asserted that "As long as large firms have both the computational resources and the

[60] See "ImageNet Large Scale Visual Recognition Challenge," at http://www.image-net.org/challenges/LSVRC/.

[61] Husanjot Chahal, Ryan Fedasiuk, and Carrick Flynn, *Messier Than Oil: Assessing Data Advantage in Military AI*, Center for Security and Emerging Technology, July 2020.

[62] Ibid.

access to proprietary datasets to combine with open data, they are likely to maintain a competitive advantage."[63] Concerns about private-sector competition and innovation constraints have been noted particularly for AI researchers and developers with limited access to data and testing and training resources, such as academic researchers, small businesses, and startups.

In response, the Select Committee on Artificial Intelligence of the National Science and Technology Council (NSTC) included "develop shared public datasets and environments for AI training and testing" as a priority area in its AI R&D Strategic Plan in 2016 and the 2019 update.[64] Additionally, the February 2019 Executive Order on *Maintaining American Leadership in Artificial Intelligence* directed the heads of all federal agencies to

> review their federal data and models to increase access and use by the greater non-federal AI research community in a manner that benefits that community, while protecting safety, security, privacy, and confidentiality. Specifically, agencies shall improve data and model inventory documentation to enable discovery and usability, and shall prioritize improvements to access and quality of AI data and models based on the AI research community's feedback.[65]

Since the national AI R&D strategic plan was first announced in 2016, numerous federal agencies have made varying degrees of progress toward collecting and sharing data. However, challenges remain, such as labeling and curating datasets so that they are useful for AI research, working with AI stakeholders to ensure that datasets and models are fit for use and are maintained as standards and norms evolve, and developing tools to verify data provenance and oversee proper use policies.

The strategic plan notes that "data alone are of little use without the ability to bring computational resources to bear on large-scale public datasets."[66] Demonstrating the intensive training needed for some systems, Facebook has described an AI experiment using billions of Instagram photos that required hundreds of graphics chips across 42 servers for almost a month.[67] An analysis by the nonprofit OpenAI found that the amount of computing power used for training certain AI systems is now rising seven times faster than it did before about 2012 (doubling every approximately 3.4 months post-2012 versus approximately 2 years pre-2012).[68] The OpenAI group recommended that policymakers consider increasing funding for academic research, as some types of AI research are becoming more computationally intensive and expensive.[69]

[63] T. Davies et al., (Eds.), "Algorithms and AI," in *State of Open Data: Histories and Horizons*, 2019, at https://www.stateofopendata.od4d.net/chapters/issues/artificial-intelligence.html.

[64] Select Committee on Artificial Intelligence, National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update*, June 2019, pp. 27-31 (hereinafter, "NSTC Select Committee on Artificial Intelligence 2019 AI R&D Strategic Plan"). See also below under "Federal Activity in AI."

[65] Executive Order 13859, "Maintaining American Leadership in Artificial Intelligence," 84 *Federal Register* 3967, February 11, 2019, at https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence.

[66] NSTC Select Committee on Artificial Intelligence 2019 AI R&D Strategic Plan, p. 28.

[67] Tom Simonite, "Your Instagram #Dogs and #Cats Are Training Facebook's AI," *Wired*, May 2, 2018, at https://www.wired.com/story/your-instagram-dogs-and-cats-are-training-facebooks-ai/.

[68] As reported by Karen Hao, "The Computing Power Needed to Train AI is Now Rising Seven Times Faster than Ever Before," *MIT Technology Review*, November 11, 2019, at https://www.technologyreview.com/2019/11/11/132004/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/.

[69] OpenAI, "AI and Compute: Addendum," *OpenAI Blog*, May 16, 2018, at https://openai.com/blog/ai-and-compute/#addendum.

Building on federal strategic planning and agency efforts to provide greater access to computational resources and high-quality data to support AI research, Congress directed the Director of the National Science Foundation in coordination with the Office of Science and Technology Policy to establish a National AI Research Resource Task Force through the National Artificial Intelligence Initiative Act of 2020.[70] The task force is to include four federal members, four members from academic institutions, and four private sector members. The task force is meant to investigate and report on the feasibility and advisability of establishing and sustaining a National Artificial Intelligence Research Resource, defined as "a system that provides researchers and students across scientific fields and disciplines with access to compute resources, co-located with publicly-available, artificial intelligence-ready government and non-government data sets and a research environment with appropriate educational tools and user support."[71]

AI Training with Small and Alternative Datasets

Some researchers have responded to the concern over limited access to big datasets for training by focusing on alternative ways to obtain or use data to reduce costs and computing power requirements. One method that has been explored is creating techniques and models that can learn from reduced amounts of data or fewer training iterations. For example, researchers at Google DeepMind created AI software that initially needs to analyze several hundred categories of images, but afterwards can learn to recognize new objects from just one picture—called "one-shot learning."[72]

Additional approaches include using alternative datasets and techniques. Some startups have reportedly created synthetic data to generate a large enough dataset for training AI models.[73] Others have demonstrated the promise of relatively unknown or novel AI techniques. For example, in recent years, some AI technologies developed by smaller AI groups have outperformed technologies from large companies such as Google and Intel in certain benchmark measures at Stanford University's DAWNBench challenge.[74] One report on this competition states,

> these metrics [such as cost and algorithm speed] help us judge whether small teams can take on the tech giants. The results don't give a straightforward answer, but they suggest that raw computing power isn't the be-all and end-all for AI success. Ingenuity in how you design your algorithms counts for at least as much. While big tech companies like Google

[70] P.L. 116-283, Division E, Section 5106.

[71] P.L. 116-283, Division E, Section 5106(g). According to information on AI.gov, information about members and meetings of the task force will be announced and posted once it is established; see https://www.ai.gov/nairrtf/#MEMBERS.

[72] Will Knight, "Machines Can Now Recognize Something After Seeing It Once," *MIT Technology Review*, November 3, 2016, at https://www.technologyreview.com/2016/11/03/6485/machines-can-now-recognize-something-after-seeing-it-once/.

[73] Tom Simonite, "Some Startups Use Fake Data to Train AI," *Wired*, April, 25, 2018, at https://www.wired.com/story/some-startups-use-fake-data-to-train-ai/.

[74] The DAWNbench challenge is an AI engineering competition in which teams and individuals from universities, governments, and industry compete to design the best algorithms, with Stanford's researchers acting as adjudicators. Each entry must meet basic accuracy standards (for example, recognizing 93% of dogs in a given dataset) and is judged on metrics training time and cost. See James Vincent, "An AI Speed Test Shows Clever Coders Can Still Beat Tech Giants Like Google and Intel," *The Verge*, May 7, 2018, at https://www.theverge.com/2018/5/7/17316010/fast-ai-speed-test-stanford-dawnbench-google-intel.

> and Intel had predictably strong showings in a number of tasks, smaller teams (and even individuals) ranked highly by using unusual and little-known techniques."[75]

AI Hardware

Hardware advances have played another key role in AI progress over the past decade, and hardware development—including AI chips and high performance computing (HPC) for AI applications—is an active research area. According to data from CB Insights, global equity funding for AI chip startups rose from just over $200 million from 13 deals in 2016 to approximately $700 million from 30 deals in 2018.[76] Companies including Nvidia, Google, Microsoft, and Facebook have been working on AI chip R&D, including developing chips designed for specialized tasks and designed to optimize energy efficiency for particular AI applications.[77]

One of the largest recent efforts in the United States to use HPC for AI applications comes from a partnership between the DOE's Oak Ridge National Laboratory (ORNL) and IBM to create the Summit supercomputer. Summit contains "AI-optimized" graphical processing units (GPUs) and has been described as "a supercomputer suited for AI."[78] The type and large number of chips allow it to run intensive ML techniques, such as DL.[79]

Sector Example: AI in Healthcare

Numerous companies and researchers have been developing and testing AI technologies for use in healthcare, for example, to detect diabetic retinopathy (an eye condition that can cause blindness in diabetic patients) and skin cancer, and to mine large quantities of medical data to derive insights.[80] Some hospitals are also experimenting with using voice recognition, and associated ML and NLP technology, to assist doctors and patients.[81]

Growth in AI and its potential healthcare applications has led to the development of various partnerships among public and private sector groups. In 2019, for example, established pharmaceutical companies partnered with startups and researchers working on AI use for drug discovery and development.[82] Federal agencies have also begun assessing the potential for AI in certain settings, such as drug discovery and clinical trials,[83] and working with

[75] James Vincent, "An AI Speed Test Shows Clever Coders Can Still Beat Tech Giants Like Google and Intel," *The Verge*, May 7, 2018, at https://www.theverge.com/2018/5/7/17316010/fast-ai-speed-test-stanford-dawnbench-google-intel.

[76] Data from CB Insights as reported in Richard Waters, "Facebook Joins Amazon and Google in AI Chip Race," *Financial Times*, February 18, 2019, at https://www.ft.com/content/1c2aab18-3337-11e9-bd3a-8b2a211d90d5.

[77] Ibid.

[78] Department of Energy, Oak Ridge National Laboratory, "Summit," at https://www.olcf.ornl.gov/summit/, and "ORNL Launches Summit Supercomputer," news release, June 8, 2018, at https://www.ornl.gov/news/ornl-launches-summit-supercomputer.

[79] Tom Simonite, "The US Again Has the World's Most Powerful Supercomputer," *Wired*, June 8, 2018, https://www.wired.com/story/the-us-again-has-worlds-most-powerful-supercomputer.

[80] Google, "Seeing Potential: How a Team at Google Is Using AI to Help Doctors Prevent Blindness in Diabetics," at https://www.google.com/about/stories/seeingpotential/; Melanie Evans and Laura Stevens, "Big Tech Expands Footprint in Health," November 27, 2018, at https://www.wsj.com/articles/amazon-starts-selling-software-to-mine-patient-health-records-1543352136; and H.A. Haenssle et al., "Man Against Machine: Diagnostic Performance of a Deep Learning Convolutional Neural Network for Dermoscopic Melanoma Recognition in Comparison to 58 Dermatologists," *Annals of Oncology*, vol. 29, no. 8 (August 1, 2018), pp. 1836-1842.

[81] Ruth Hailu, "5 Burning Questions About Deploying Voice Recognition Technology in Health Care," *STAT News*, July 10, 2019, at https://www.statnews.com/2019/07/10/5-questions-voice-recognition-technology/.

[82] Robert Langreth, "AI Drug Hunters Could Give Big Pharma a Run for Its Money, *Bloomberg*, July 15, 2019, at https://www.bloomberg.com/news/features/2019-07-15/google-ai-could-challenge-big-pharma-in-drug-discovery.

[83] Government Accountability Office, *Artificial Intelligence in Health Care: Benefits and Challenges of Machine*

the private sector to evaluate the use of AI systems. For example, a partnership between the Department of Veterans Affairs and DeepMind has worked to identify risk factors for patient deterioration during hospitalization in an effort to develop early interventions and improve care.[84] Further, the Food and Drug Administration has been developing a framework for regulating AI- and ML-based software as a medical device and addressing subsequent modifications to such software.[85]

While there are many encouraging developments for using AI technologies in healthcare, stakeholders have remarked on the slow progress in using AI broadly within healthcare settings, and various challenges and questions remain. Researchers and clinicians have raised questions about the accuracy, security, and privacy of these technologies; the availability of sufficient health data on which to train systems; medical liability in the event of adverse outcomes; patient access and receptivity; and the adequacy of current user consent processes.[86] A 2019 literature review and meta-analysis of the performance of DL systems compared to medical professionals in detecting disease from medical imaging concluded that few of the 82 identified studies presented externally validated results and "poor reporting is prevalent in deep learning studies, which limits reliable interpretation of the reported diagnostic accuracy," concluding that new reporting standards could improve future studies.[87]

Federal Activity in AI

In recent years, the federal government—including the White House, federal agencies, and Congress—has increasingly supported and conducted AI R&D, invested in AI technologies, and worked to address issues with AI development and use. AI has been of interest to Congress since at least the 1980s and congressional AI activities, including legislation and oversight hearings, increased in the 115th and 116th Congresses.[88] This section of the report focuses on selected federal activities during the Administrations of Donald J. Trump and Barack Obama and in the 115th and 116th Congresses.

Executive Branch

The Trump and Obama Administrations took a variety of actions related to AI, by establishing initiatives through executive order, forming committees, and releasing reports. Further, in accordance with the National Artificial Intelligence Initiative Act of 2020 (P.L. 116-283, Division E, as described in the "Legislation" section), the Office of Science and Technology Policy

Learning in Drug Development, GAO-20-215SP, January 21, 2020, at https://www.gao.gov/products/GAO-20-215SP.

[84] Department of Veterans Affairs, Office of Public and Intergovernmental Affairs, "VA Partners With DeepMind to Build Machine Learning Tools to Identify Health Risks for Veterans," February 21, 2018, at https://www.va.gov/opa/pressrel/pressrelease.cfm?id=4013.

[85] See U.S. Food and Drug Administration, "Artificial Intelligence and Machine Learning in Software as a Medical Device," at https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.

[86] Ruth Hailu, "5 Burning Questions About Deploying Voice Recognition Technology in Health Care," *STAT News*, July 10, 2019, at https://www.statnews.com/2019/07/10/5-questions-voice-recognition-technology/; and Lauren Joseph, "5 Burning Questions About Using Artificial Intelligence to Prevent Blindness," *STAT News*, July 17, 2019, at https://www.statnews.com/2019/07/17/artificial-intelligence-to-prevent-blindness/.

[87] Xiaoxuan Liu et al., "A Comparison of Deep Learning Performance Against Health-Care Professionals in Detecting Diseases from Medical Imaging: A Systematic Review and Meta-Analysis," *The Lancet*, vol. 1, no. 6 (October 1, 2019), pp. E271-E297.

[88] For example, see U.S. Congress, Subcommittee on Investigations and Oversight, Committee on Science and Technology, U.S. House of Representatives, *Robotics*, 97th Congress, 2nd sess., June 2 and 23, 1982 (Washington, DC: GPO, 1983).

(OSTP) launched the National AI Initiative Office (NAIIO) on January 12, 2021, to coordinate and support the National AI Initiative (the act is further described in the "Legislation" section).[89]

Executive Orders on AI

In February 2019, President Trump released an executive order establishing the American AI Initiative (E.O. 13859).[90] In addition to promoting AI R&D investment and coordination, objectives of the E.O. include making federal data, models, and computing resources available for AI development, reducing barriers to the use of AI technologies, developing technical and international standards around AI innovation, preparing an action plan around AI and national security concerns, and training the workforce to develop and use AI.

In December 2020, President Trump released an executive order promoting the use of trustworthy AI in the federal government (E.O. 13960).[91] The E.O. establishes a common set of principles for the design, development, acquisition, and use of AI in the federal government to foster public trust and confidence, and directs the Office of Management and Budget (OMB) to develop policy guidance for implementing the principles across agencies. The E.O. further includes direction to federal agencies (1) to provide annual, publicly-available inventories of non-classified, non-sensitive use cases of AI, and (2) to undertake activities to expand the number of AI experts at federal agencies, including through creating an AI track within the Presidential Innovation Fellows program and by assessing potential expansion of federal rotational programs.

National Science and Technology Council Committees

The National Science and Technology Council (NSTC) convenes federal science and technology leaders as a primary means within the executive branch to coordinate science and technology policies across federal agencies.[92] The Trump Administration established a new committee and expanded on committees and working groups established by the Obama Administration, with the following NSTC bodies coordinating cross-agency efforts in AI and ML:

- The Select Committee on Artificial Intelligence was established in May 2018 and re-chartered on January 5, 2021 "in accordance with the National Artificial Intelligence Act of 2020 … with a broader scope and membership." The Committee is comprised of heads of agencies and advises the White House on interagency AI R&D priorities; provides a formal mechanism for interagency policy coordination and the development of federal AI activities; and addresses national and international AI policy matters.[93]

[89] See information provided by National Artificial Intelligence Initiative Office, "NAIIO—National Artificial Intelligence Initiative Office," at https://www.ai.gov/about/#NAIIO_National_Artificial_Intelligence_Initiative_Office. The "AI.gov" website was originally launched by the Trump Administration; a new version of the website was launched by the Biden Administration on May 5, 2021.

[90] Executive Order 13859, "Maintaining American Leadership in Artificial Intelligence," 84 *Federal Register* 3967, February 11, 2019.

[91] Executive Order 13960, "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government," 85 *Federal Register* 78939, December 8, 2020, at https://www.federalregister.gov/d/2020-27065.

[92] For additional information on the NSTC, see CRS Report R43935, *Office of Science and Technology Policy (OSTP): History and Overview*, by John F. Sargent Jr. and Dana A. Shea.

[93] For additional information on the NSTC Select Committee on Artificial Intelligence, see the January 5, 2021 charter, at https://trumpwhitehouse.archives.gov/wp-content/uploads/2021/01/Charter-Select-Committee-on-AI-Jan-2021-posted.pdf.

- The ML and AI (MLAI) Subcommittee is the operations and implementation arm of the Select Committee on Artificial Intelligence and includes federal employees with budgetary decisionmaking responsibilities to help focus priorities for AI investments through agency programs.
- The AI Interagency Working Group is a community of practice,[94] taking on tasks that require deep expert knowledge and producing products such as the AI R&D Strategic Plan and its updates.[95]

Select AI Reports and Documents

As federal government interest and engagement in AI has grown, the executive branch has included a focus on AI in a variety of strategic plans, reports, and memoranda, including the following.

- The NSTC first released the *National AI Research and Development Strategic Plan* in 2016 with seven strategic priorities.[96] In September 2018, NITRD's National Coordination Office requested input from the public on whether and how the plan should be revised and improved.[97] In response, various industry groups requested more detail on federal priorities in AI R&D—including on specific challenges, applications, ways to incorporate private sector participation, and goals for investments from both technical and social impact perspectives. Some groups also asserted a need to align federal plans for enabling technologies such as 5G and quantum computing with the AI strategy.[98] In June 2019, NSTC released an updated plan with eight strategic priorities, the last of which was new: (1) make long-term investments in AI research; (2) develop effective methods for human-AI collaboration; (3) understand and address ethical, legal, and societal implications of AI; (4) ensure the safety and security of AI systems; (5) develop shared public datasets and environments for AI training and testing; (6) measure and evaluate AI technologies through standards and benchmarks; (7) better understand the national AI R&D workforce needs; and (8) expand public-private partnerships to accelerate advances in AI.[99]
- In August 2019, in response to E.O. 13859, NIST released the *Plan for Federal Engagement in Developing Technical Standards and Related Tools in AI*. NIST

[94] A community of practice is generally a group of professionals who are active in, or interested in, a particular craft or profession. For example, the General Service Administration (GSA) also leads an AI community of practice to "bring together federal employees who are active in, or interested in, AI policy technology, standards, and programs to facilitate the sharing of best practices, use cases, and lessons learned; and [to] advance and share tools, playbooks success stories with a community of interested professionals." Steven Babitch, "GSA Launches Artificial Intelligence Community of Practice," *GSA Blog*, November 5, 2019, at https://www.gsa.gov/blog/2019/11/05/gsa-launches-artificial-intelligence-community-of-practice.

[95] Overviews of the activities of each body include descriptions provided during a telephone conversation between CRS and Dr. Lynne Parker, Deputy Chief Technology Officer of the United States, March 2019.

[96] National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, *The National Artificial Intelligence Research and Development Strategic Plan*, October 2016.

[97] NITRD National Coordination Office, "Request for Information on Update to the 2016 National Artificial Intelligence Research and Development Strategic Plan," 83 *Federal Register* 48655, September 26, 2018.

[98] MeriTalk, "Industry Wants More Detail on AI R&D Plan," December 21, 2018, at https://www.meritalk.com/articles/industry-wants-more-detail-on-ai-rd-plan/.

[99] NSTC Select Committee on Artificial Intelligence 2019 AI R&D Strategic Plan.

noted that the plan was prepared with broad public and private sector input. It includes recommendations for federal government activities to engage in deep, long-term AI standards development "to speed the pace of reliable, robust, and trustworthy AI technology development."[100]

- In August 2020, OMB and OSTP provided their annual memorandum to the heads of federal R&D agencies laying out the Administration's R&D budget priorities for FY2022. The memorandum stated that industries of the future—including AI—remained a top R&D priority for the Administration, as in prior years.[101]
- In November 2020, OMB released a memorandum to the heads of federal agencies providing guidance for the regulation of AI. The purpose of the memo was to guide regulatory and non-regulatory oversight of AI applications developed and deployed outside of the federal government. It lays out 10 principles for the stewardship of AI applications, including topics such as risk assessment, fairness and nondiscrimination, disclosure and transparency, and interagency coordination. It further touches on reducing barriers to the deployment and use of AI, including increasing access to government data, communicating benefits and risks to the public, engaging in the development and use of voluntary consensus standards, and engaging in international regulatory cooperation efforts. Agency plans to conform to the memorandum are due on May 17, 2021, and must include any statutory authorities governing agency regulation of AI applications, information collections on AI from regulated entities, regulatory barriers to AI applications, and any planned or considered regulatory actions on AI.[102]

In addition to the initial National AI R&D Strategic Plan, two other background documents on AI were also prepared in 2016 by the NSTC and other offices in the Executive Office of the President. These reports were *Preparing for the Future of Artificial Intelligence*, and *Artificial Intelligence, Automation, and the Economy*.[103]

Federal Agency Activities

Engagement on AI varies across agencies and may include examining and adopting AI technologies for internal agency use, holding hearings to examine issues surrounding the development and use of AI,[104] conducting AI R&D in-house (intramural R&D), and funding AI

[100] National Institute of Standards and Technology, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, August 9, 2019, pp. 3-6.

[101] Office of Management and Budget and Office of Science and Technology Policy, "Memorandum for the Heads of Executive Departments and Agencies: Fiscal Year (FY) 2022 Administration Research and Development Budget Priorities and Cross-cutting Actions" August 14, 2020, at https://www.whitehouse.gov/wp-content/uploads/2020/08/M-20-29.pdf.

[102] Russell Vought, Director of the Office of Management and Budget, "Guidance for Regulation of Artificial Intelligence Applications," Memorandum for the heads of executive departments and agencies, November 17, 2020, at https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf.

[103] Available at https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf; and https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF.

[104] For example, the Federal Trade Commission (FTC) held a hearing in November 2018 focused on consumer welfare implications associated with the use of algorithmic decision tools, AI, and predictive analytics; see https://www.ftc.gov/news-events/events-calendar/ftc-hearing-7-competition-consumer-protection-21st-century.

R&D by outside groups (extramural R&D), including at institutions of higher education (IHEs), nonprofits, and industry. E.O. 18539 directed federal R&D agencies to "promote sustained investment in AI R&D in collaboration with industry, academia, international partners and allies, and other non-Federal entities" and the heads of those agencies to consider AI as an R&D priority when preparing their budget requests to Congress. E.O. 13960 highlighted a range of ways that federal agencies are already employing AI, including identifying information security threats, facilitating review of large datasets, streamlining processes for grant applications, modeling weather patterns, and facilitating predictive maintenance.

Although there are numerous examples of federal agencies using AI in-house, there is currently no comprehensive database of AI projects within agencies, though some recent efforts have attempted to better compile such information.[105] The General Services Administration (GSA) has reportedly been working to catalogue some use cases of AI across the federal government.[106] Additionally, the Administrative Conference of the United States (ACUS) commissioned a study, completed in February 2020, "to map how federal agencies are currently using AI to make and support decisions."[107] Among 142 federal agencies, the study authors identified use cases—defined as "instance[s] in which an agency had considered using or had already deployed AI/ML technology to carry out a core function"—in 64 (45%) agencies, based on searches of publicly available information.[108] Of the 157 use cases, the authors noted that 84 (53%) were built in-house, rather than being procured through private contracting or non-commercial collaboration (e.g., with an academic laboratory or through a public-facing competition).[109] Building on this initial study, E.O. 13960 requires federal agencies to create publicly available inventories of use cases of AI, based on common criteria, format, and inventory mechanisms created by the Federal Chief Information Officers Council.

Some examples of federal agencies using AI in-house include the following:

- The Department of Health and Human Services used AI and NLP technologies to identify incorrect citations and outdated regulations in the Code of Federal Regulations as part of a "department-wide regulatory clean-up initiative."[110]
- NASA launched RPA pilot projects in accounts payable and receivable, IT spending, and human resources. The projects appeared to work well—in the human resources application, for example, 86% of transactions were completed without human intervention—and are being rolled out across the organization. NASA reportedly moved forward with implementing more RPA bots, some with higher levels of intelligence.[111]

[105] CRS communications with the Office of Science and Technology Policy, February 2020; and David Freeman Engstrom et al., *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*, Report delivered to the Administrative Conference of the United States, February 2020.

[106] CRS communications with the Office of Science and Technology Policy, February 2020.

[107] See Administrative Conference of the United States, Office of the Chairman Projects, "Artificial Intelligence in Federal Agencies," February 2020, at https://www.acus.gov/research-projects/artificial-intelligence-federal-agencies (hereinafter, "ACUS report 2020").

[108] Ibid, pp. 15-18. The authors limited the included agencies to those with over 400 employees and excluded active military and intelligence-related organizations.

[109] Ibid. p. 18.

[110] Department of Health and Human Services, "HHS Launches First-of-Its-Kind Regulatory Clean-Up Initiative Utilizing AI," November 17, 2020, at https://www.hhs.gov/about/news/2020/11/17/hhs-launches-first-its-kind-regulatory-clean-up-initiative-utilizing-ai.html.

[111] Thomas H. Davenport and Rajeev Ronanki, "Artificial Intelligence for the Real World," *Harvard Business Review*,

- The National Oceanic and Atmospheric Administration (NOAA) has developed an AI strategy to "expand the application of [AI] in every NOAA mission area by improving the efficiency, effectiveness, and coordination of AI development and usage across the agency."[112]
- The Social Security Administration has used AI/ML in its adjudication work to address challenges from high caseloads and in ensuring accuracy and consistency of decisionmaking, which have reportedly persisted through decades of quality improvement efforts.[113]

Considerations for agency adoption of AI mirror private sector considerations—namely, how can AI be used as a tool to advance process automation, provide insight into data analyses, and improve services (i.e., improve timeliness and enhance citizen interactions with federal agencies, such as through the use of chatbots). Technology leaders in federal agencies, industry, and academia have argued that the initial implementation of AI technologies should be evaluated in terms of challenges and opportunities associated with an agency's current data collection, management, and analysis processes, rather than the capabilities of AI systems themselves.[114] Additional considerations include how to evaluate and acquire AI systems.

To further guide agencies, E.O. 13960 provides broad principles for federal design, development, acquisition, and use of AI, including that AI systems should be (1) lawful and respectful of the nation's values; (2) purposeful and performance-driven; (3) accurate, reliable, and effective; (4) safe, secure and resilient; (5) understandable; (6) responsible and traceable; (7) regularly monitored; (8) transparent; and (9) accountable. Given that OMB is tasked with developing, by June 2021, a roadmap for policy guidance to better support federal government use of AI, more concrete plans and actions may be specified across agencies.

The National Science Foundation (NSF) has been a primary nondefense source of federal extramural support for AI R&D for decades and currently "supports fundamental research, education and workforce development, and advanced, scalable computing resources that collectively enhance fundamental research in AI."[115] Fundamental AI research areas include how computer systems represent knowledge; learn; process spoken and written language; and solve problems, as well as the impacts of AI on continuing education and adult retraining.[116] Additional federal agency activities in AI R&D include:

- NIST engaged in national and international AI standards development activities;
- DARPA launched the AI Next campaign, focused on "improving the robustness and reliability of AI systems; enhancing the security and resiliency of machine

January-February 2018, pp. 108-116, at https://hbr.org/2018/01/artificial-intelligence-for-the-real-world.

112 National Oceanic and Atmospheric Administration, *NOAA Artificial Intelligence Strategy: Analytics for Next-Generation Earth Science*, February 2020, at https://nrc.noaa.gov/LinkClick.aspx?fileticket=0I2p2-Gu3rA%3d&tabid=91&portalid=0.

113 Administrative Conference of the United States, "Artificial Intelligence in Federal Agencies," February 2020, pp. 38-39.

114 See remarks by Stephen Dennis, Director of the Data Analytics Engine, Science and Technology Directorate, Department of Homeland Security at the FCW Workshop, "Artificial Intelligence: Moving from Vision to Implementation," March 13, 2018; and Davenport and Ronanki, 2018.

115 Information about NSF support for AI research and workforce programs and interagency work can be found at "Artificial Intelligence at NSF," at https://www.nsf.gov/cise/ai.jsp.

116 NSF, "Statement on Artificial Intelligence for American Industry," press statement 18-005, May 10, 2018, at https://www.nsf.gov/news/news_summ.jsp?cntn_id=245418.

learning and AI technologies; reducing power, data, and performance inefficiencies; and pioneering the next generation of AI algorithms and applications, such as 'explainability' and common sense reasoning";[117]

- DOE established the Artificial Intelligence and Technology Office to "accelerate the delivery of AI-enabled capabilities, scale the department-wide development and impact of AI, and synchronize AI activities to advance the agency's core missions, expand partnerships, and support American AI Leadership";[118]
- The Department of Veteran's Affairs (VA) established a National Artificial Intelligence Institute (NAII) to develop AI R&D capabilities in the VA;[119] and
- The National Institute for Justice—the research wing of the Department of Justice—supported research on "crime-fighting AI" which "it believes could be used to fight human trafficking, illegal border crossings, drug trafficking, and child pornography" by helping investigators sort through data.[120]

Congress

The 115th and 116th Congresses focused on AI more frequently and explicitly than previous Congresses, in terms of enacted and introduced legislation and hearings. Additionally, bipartisan AI caucuses were launched in the House and the Senate.[121] The AI Index group used data from McKinsey & Company to assess mentions of AI in Congress based on the *Congressional Record*. The analysis found, after a maximum of 9 mentions in any year from 2011 through 2016, mentions increased each year throughout the 115th and 116th Congresses, with 129 mentions reported in 2020 (**Figure 3**).[122] This section of the report provides a brief summary of legislative activities in the 116th and 117th Congresses, including descriptions of laws and selected bills that focused on, or included specific provisions focused on AI and ML, as well as hearings from the 115th-117th Congresses (as of the date of this report).[123]

[117] DARPA, "DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies," September 7, 2018, at https://www.darpa.mil/news-events/2018-09-07.

[118] See U.S. Department of Energy, Artificial Intelligence and Technology Office, at https://www.energy.gov/science-innovation/artificial-intelligence-and-technology-office.

[119] See U.S. Department of Veterans Affairs, Office of Research and Development, "National Artificial Intelligence Institute (NAII)," at https://www.research.va.gov/naii/.

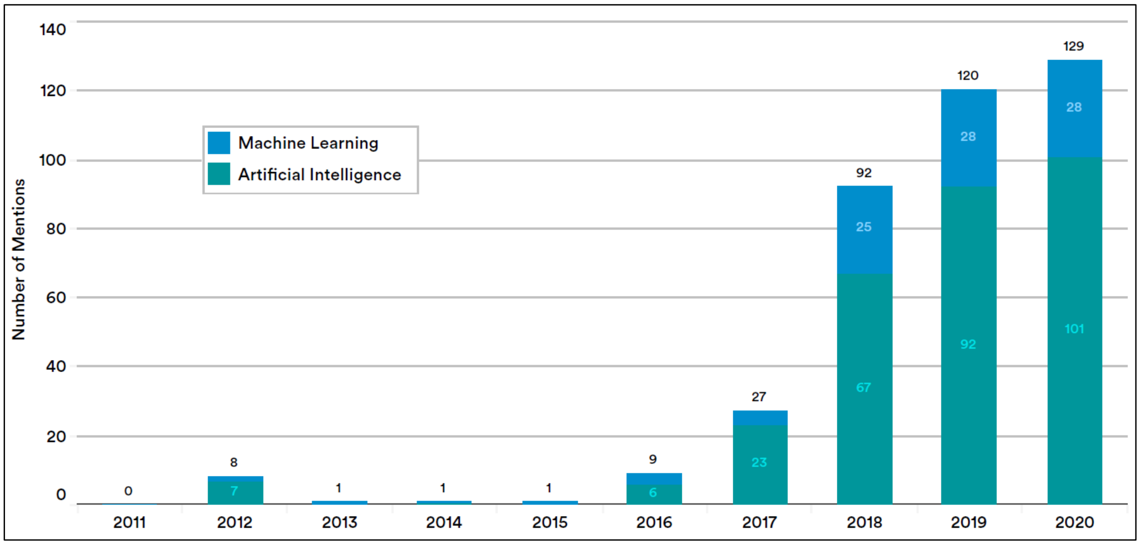
[120] Kate Conger, "Justice Department Drops $2 Million to Research Crime-Fighting AI," *Gizmodo*, February 27, 2018; and DOJ's solicitation for the program can be found at https://nij.gov/funding/Documents/solicitations/NIJ-2018-14000.pdf.

[121] The House Congressional AI Caucus was originally launched in 2015; see https://artificialintelligencecaucus-olson.house.gov/. The Senate AI Caucus was announced on March 13, 2019; see announcements from the caucus co-chairs at https://www.portman.senate.gov/public/index.cfm/2019/3/portman-heinrich-launch-bipartisan-artificial-intelligence-caucus, and https://www.heinrich.senate.gov/press-releases/heinrich-portman-launch-bipartisan-artificial-intelligence-caucus.

[122] AI Index 2021 Annual Report, p. 172.

[123] Additional bills mentioned AI or ML without including specific provisions related to the technologies. For example, the Developing Innovation and Growing the Internet of Things Act (S. 1611, 116th Congress) stated in the findings that "the Internet of things will … play a key role in developing artificial intelligence and advanced computing capabilities," but AI was not included anywhere else in the bill. Such bills are not discussed in this section.

Figure 3. Mentions of Artificial Intelligence and Machine Learning in the *Congressional Record*, 2011-2020



Source: AI Index Steering Committee, *The AI Index 2021 Annual Report*, Human-Centered AI Institute, Stanford University, Stanford, CA, March 2021, pp. 171-172; data from the McKinsey Global Institute, 2020.

Notes: Per the AI Index 2021 Annual Report, each count indicates that AI or ML was mentioned during a particular event contained in the *Congressional Record*, including the reading of a bill. If a speaker or member mentioned AI or ML multiple times within remarks, or multiple speakers mentioned AI or ML within the same event, it appears only once as a result. Counts for AI and ML are separate, as they were conducted in separate searches. Mentions of the abbreviations "AI" or "ML" are not included. Additional information about the search methodology is included in the AI Index 2021 Annual Report appendix, p. 216.

Legislation

As of the date of this report, multiple bills introduced in the 117th Congress have included language about AI, either as a focus of the bill or in a specific provision, though no legislation has been enacted. Some bills have included AI as one of multiple key technology areas important for U.S. competitiveness.[124] Other bills have focused on federal AI expertise; addressed potential bias in automated decision systems that may use AI; or included AI as a technology with potential applications in healthcare.[125]

At least four laws enacted in the 116th Congress focused on AI or included AI-focused provisions. The FY2021 NDAA included multiple sections related to Department of Defense (DOD) AI activities in R&D, acquisitions, and workforce expansion and training. These sections built on prior direction in the FY2020 NDAA, which included provisions related to recruiting expertise at the DOD Joint Artificial Intelligence Center (JAIC); establishing DOD processes to update policies on emerging technologies, including AI; extending authorization for the National Security Commission on Artificial Intelligence; and requiring an analysis of major initiatives of the intelligence community in AI and ML.

[124] For the 117th Congress, see, for example, the Endless Frontier Act (S. 1260) and the Strategic Competition Act of 2021 (S. 1169), the STRATEGIC Act (S. 687), and the Democracy Technology Partnership Act (S. 604).

[125] For the 117th Congress, see, for example, "A bill to establish a Federal artificial intelligence scholarship-for-service program" (S. 1257), the Unemployment Insurance Technology Modernization Act of 2021 (S. 490); the Black Maternal Health Momnibus Act of 2021 (S. 346 and H.R. 959); and the Tech to Save Moms Act (H.R. 937).

Further, the FY2021 NDAA incorporated the expansive National Artificial Intelligence Act of 2020 (Division E), which included sections related to

- Codifying the establishment of an American AI Initiative (Section 5101);
- Establishing the National AI Initiative Office to support federal AI activities, including technical, programmatic, and administrative support for activities of the AI Initiative, as specified (Section 5102);
- Establishing an Interagency Committee at OSTP to coordinate federal programs and activities in support of the AI Initiative, including developing periodic strategic plans for AI (Section 5103);[126]
- Establishing a National AI Advisory Committee with representatives from academic institutions, companies, nonprofit and civil society entities, and federal laboratories to provide to the President and the AI Initiative Office "advice and information on science and technology research, development, ethics, standards, education, technology transfer, commercial application, security, and economic competitiveness" related to AI (Section 5104(a));
- Establishing as part of the National AI Advisory Committee a Subcommittee on AI and Law Enforcement to provide advice on bias, data security, adoptability, and legal standards (Section 5104(e));
- Directing NSF to contract with the National Academies of Sciences, Engineering, and Medicine to conduct a study on the current and future impact of AI on the U.S. workforce across sectors (Section 5105);
- Establishing a task force to investigate the feasibility of, and plan for, a National AI Research Resource, defined as "a system that provides researchers and students across scientific fields and disciplines with access to compute resources, co-located with publicly-available, AI-ready government and non-government data sets and a research environment with appropriate educational tools and user support" (Section 5106);
- Directing NSF to establish a program to support a network of National AI Research Institutes, which shall be public-private partnerships that focus on a particular economic or social sector and associated ethical, societal, safety, and security implications, or a cross-cutting challenge for AI systems, with the potential to create or enhance innovation ecosystems and support interdisciplinary R&D, education, and workforce development in AI (Section 5201);[127]
- Directing NIST to support AI standards development, develop a risk management framework for trustworthy AI systems, and develop best practices for documenting and sharing data sets used to train AI systems (Section 5301);
- Directing the NOAA to establish a Center for AI (Section 5303);

[126] This section effectively expanded on and codified the NSTC Select Committee on Artificial Intelligence that was established in the Trump Administration.

[127] NSF began funding National AI Research Institutes in FY2020 in a joint effort with the U.S. Department of Agriculture National Institute of Food and Agriculture, the Department of Homeland Security Science and Technology Directorate, and the Department of Transportation Federal Highway Administration; see NSF's program description page at https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505686.

- Directing NSF to fund research and education activities in AI and related fields (Section 5401); and
- Directing the DOE to carry out a cross-cutting R&D program to advance AI tools, systems, capabilities, and workforce needs and to improve the reliability of AI methods and solutions relevant to DOE's mission (Section 5501)."

The Consolidated Appropriations Act, 2021 (P.L. 116-260) included the AI in Government Act of 2020 (Division U, Title I), which created within the General Services Administration (GSA) an AI Center of Excellence (CoE) to facilitate the adoption of AI technologies in the federal government.[128] The AI CoE is further required, among other activities, to collect, aggregate, and publish on a publicly available website information regarding federal programs, pilots, and other initiatives; and to advise federal agencies on the acquisition and use of AI through technical insight and expertise. The act required OMB to issue a memorandum to federal agencies regarding the development of AI policies, approaches for removing barriers to using AI technologies, and best practices for identifying, assessing, and mitigating any discriminatory impact or bias and any unintended consequences of using AI. Additionally, the act required the Office of Personnel Management to establish or update an occupational job series to include positions with primary duties in AI and to estimate current and future numbers of federal employment positions related to AI at each agency.

The Further Consolidated Appropriations Act, 2020 (P.L. 116-94) included a provision amending the Export-Import Bank Act of 1945 to establish a Program on China and Transformational Exports (Section 402). This program is directed to support the extension of loans, guarantees, and insurance that aim to "advance the comparative leadership of the United States with respect to the People's Republic of China, or support United States innovation, employment, and technological standards, through direct exports in" artificial intelligence, among other areas.

The Identifying Outputs of Generative Adversarial Networks Act (P.L. 116-258) directed NSF and NIST to support research on generative adversarial networks, including research on manipulated or synthesized content and information authenticity and the development of measurements and standards necessary to accelerate the development of technical tools to examine the function and outputs of GANs.

Multiple additional bills introduced in the 116th Congress address AI applications, such as facial recognition and deepfakes,[129] and areas in which AI is deployed, including law enforcement and criminal justice, healthcare, energy efficiency, natural resources, and defense and national security.[130] Some of these bills are focused on AI, while others include AI-specific provisions as part of a broader focus.

[128] The act codified the GSA AI Center of Excellence that was launched in 2019; see https://www.ai.gov/legislation-and-executive-orders/.

[129] For the 116th Congress, see, for example, the Ethical Use of Facial Recognition Act (S. 3284); the Facial Recognition Technology Warrant Act of 2019 (S. 2878); the Facial, Analysis, Comparison, and Evaluation (FACE) Protection Act of 2019 (H.R. 4021); the Commercial Facial Recognition Privacy Act of 2019 (S. 847); the Deepfakes Report Act (H.R. 3600 and S. 2065); and the Deep Fake Detection Prize Competition Act (H.R. 5532).

[130] For the 116th Congress, see, for example, the Advancing Innovation to Assist Law Enforcement Act (H.R. 2613); the Black Maternal Health Momnibus Act of 2020 (S. 3424, H.R. 6142); the Department of Energy Veterans' Health Initiative Act (S. 143 and H.R. 617); the Securing American Leadership in Science and Technology Act of 2020 (H.R. 5685); and the BLUE GLOBE Act (H.R. 3548), in addition to the aforementioned provisions in the National Defense Authorization Acts in FY2019 (P.L. 115-232) and FY2020 (P.L. 116-92).

Hearings

Various committees in both the House of Representatives and the Senate held hearings focused on issues in AI and ML during the 115th, 116th, and 117th Congresses. Given its many, wide-ranging applications, the topic of AI has arisen as a consideration during numerous hearings. Hearing subjects with an explicit focus on AI and ML have ranged from broad considerations of AI and ML technologies and policies, including societal and ethical issues,[131] international research and competition,[132] and national security,[133] to more focused topics, such as use by the federal government,[134] potential impact to the U.S. workforce,[135] and consequences for human rights.[136] Hearings have also focused on specific AI applications, such as facial recognition and deepfakes,[137] and contact tracing for COVID-19 cases,[138] as well as use areas, such as financial services[139] and counterterrorism.[140]

Additionally, in the 115th Congress, the House Committee on Oversight and Government Reform held a series of three hearings focusing on AI: "Game Changers: Artificial Intelligence Part 1" on February 14, 2018; "Game Changers: Artificial Intelligence Part II, Artificial Intelligence and the Federal Government" on March 7, 2018; and "Game Changers: Artificial Intelligence and Public Policy" on April 18, 2018. Subsequently, the chairman and ranking member of the Subcommittee on Information Technology released a white paper summarizing lessons learned from the hearings

[131] U.S. Congress, House Committee on Science, Space, and Technology, *Artificial Intelligence: Societal and Ethical Implications*, 116th Cong., 1st sess., June 26, 2019.

[132] U.S. Congress, Joint U.S.-China Economic and Security Review Commission, *Hearing on Technology, Trade, and Military-Civil Fusion: China's Pursuit of Artificial Intelligence, New Materials, and New Energy*, 116th Cong., 1st sess., June 7, 2019.

[133] For example, U.S. Congress, Senate Committee on Armed Services, *Emerging Technologies and Their Impact on National Security*, 117th Cong., 1st sess., February 23, 2021, at https://www.armed-services.senate.gov/hearings/21-02-23-emerging-technologies-and-their-impact-on-national-security.

[134] For example, U.S. Congress, House Committee on Science, Space, and Technology, Subcommittee on Research and Technology and Subcommittee on Energy, *Artificial Intelligence: With Great Power Comes Great Responsibility*, 115th Cong., 2nd sess., June 26, 2018; and U.S. Congress, Senate Committee on Armed Services, Subcommittee on Emerging Threats and Capabilities, *Artificial Intelligence Initiatives Within the Department of Defense*, 116th Cong., 1st sess., March 12, 2019.

[135] U.S. Congress, House Committee on the Budget, *Machines, Artificial Intelligence, and the Workforce: Recovering and Readying Our Economy for the Future*, 116th Cong., 2nd sess., September 10, 2020; and U.S. Congress, House Committee on Science, Space, and Technology, Subcommittee on Research and Technology, *Artificial Intelligence and the Future of Work*, 116th Cong., 1st sess., September 24, 2019.

[136] U.S. Congress, House of Representatives, Tom Latos Human Rights Commission, *Artificial Intelligence: The Consequences for Human Rights*, 115th Cong., 2nd sess., May 22, 2018.

[137] U.S. Congress, House Permanent Select Committee on Intelligence, *National Security Challenges of Artificial Intelligence, Manipulated Media, and "Deepfakes,"* 116th Cong., 1st sess., June 13, 2019.

[138] U.S. Congress, House Committee on Financial Services, Task Force on Artificial Intelligence, *Virtual Hearing—Exposure Notification and Contact Tracing: How AI Helps Localities Reopen Safely and Researchers Find a Cure*, 116th Cong., 2nd sess., July 8, 2020, at https://financialservices.house.gov/calendar/eventsingle.aspx?EventID=406731.

[139] The House Committee on Financial Services established a Task Force on AI in May 2019, to examine issues including AI in financial services regulation, risk management, digital identification and combatting fraud, and reducing AI bias; see for example, U.S. Congress, House Committee on Financial Services, Task Force on AI, *Equitable Algorithms: Examining Ways to Reduce AI Bias in Financial Services*, 116th Cong., 2nd sess., Feb. 12, 2020, at https://financialservices.house.gov/calendar/eventsingle.aspx?EventID=406120; and U.S. Congress, House Committee on Financial Services, Task Force on AI, *Equitable Algorithms: How Human-Centered AI Can Address Systemic Racism and Racial Justice in Housing and Financial Services*, 117th Cong., 1st sess., May 7, 2021.

[140] U.S. Congress, House Committee on Homeland Security, Subcommittee on Intelligence and Counterterrorism, *Artificial Intelligence and Counterterrorism: Possibilities and Limitations*, 116th Cong., 1st sess., June 25, 2019.

and related oversight activities, as well as recommendations for the federal government in moving forward on AI. Broadly, the recommendations included increased engagement on AI by Congress and the Administration, including increased federal R&D funding; increased stakeholder engagement in developing strategies to improve worker education, training, and reskilling; agency reviews of federal privacy laws and regulatory frameworks; and assurance that AI systems are "accountable and inspectable" when agencies use them for decisionmaking about people.[141]

Selected Issues for Congressional Consideration

Though specific AI technologies and application areas each have their own benefits, challenges, and policy issues, this section of the report will focus on some broad, crosscutting issues, with application-specific examples.

The broad potential benefits of AI technologies include opportunities for speed of data analysis and insights into big datasets, such as identification of patterns; augmentation of human decisionmaking; performance optimization for complex tasks and systems; and improved safety for people in dangerous occupations. For example, AI systems can improve facilities operations and efficiency, providing cost savings. In one application of such benefits, DeepMind reported applying ML to Google data centers to make recommendations to reduce the amount of energy used for cooling by up to 40%, subsequently moving to autonomous operations.[142]

At the same time, there are challenges and pitfalls associated with deployment and use of AI systems. For example, AI systems may perpetuate or amplify bias (as described in the "Ethics, Bias, Fairness, and Transparency" section) and may not yet be able to fully explain their decisionmaking (sometimes referred to as the "black box" problem), which can be particularly problematic in high-stakes situations, for example when they inform health and safety decisions. To train and evaluate complex AI systems, researchers and developers may need large datasets that are not widely accessible. Further, stakeholders have questioned the adequacy of public and private sector workforces to develop and work with AI, as well as the adequacy of current laws and regulations in dealing with societal and ethical issues that may arise.

In response to such overarching considerations, Congress might weigh the potential benefits of AI, such as increasing human safety, health, and productivity, with potential consequences, intended or otherwise, including job displacement and biases in algorithmic decisionmaking, when considering potential AI funding, policies, and regulation.

The passage of the National Artificial Intelligence Initiative Act of 2020 included provisions that directed federal government-wide activities and touched on many of the AI-associated issues raised in this report. Subsequently, Congress may decide that no additional legislative action is currently necessary, instead focusing in the near term on oversight of the implementation and effectiveness of the activities and programs directed by the act. This, along with activities begun in response to the aforementioned E.O.s, may provide better data and information for developing future legislation and congressional activities. Alternatively, given the rapid development of AI

[141] Rep. Will Hurd and Rep. Robin Kelly, "Rise of the Machines: Artificial Intelligence and Its Growing Impact on U.S. Policy," Subcommittee on Information Technology, Committee on Oversight and Government Reform, U.S. House of Representatives, September 2018.

[142] DeepMind, "DeepMind AI Reduces Google Data Centre Cooling Bill by 40%," July 20, 2016, at https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40; and Google, "Safety-First AI for Autonomous Data Center Cooling and Industrial Control," August 17, 2018, at https://www.blog.google/inside-google/infrastructure/safety-first-ai-autonomous-data-center-cooling-and-industrial-control/.

technologies and the wide range of sectors in which AI is deployed, Congress may decide that more actions are necessary to begin addressing issues surrounding AI use. Several major issues associated with the further development and use of AI and policy questions that Congress might consider are discussed below.

Implications for the U.S. Workforce

Concerns about job losses resulting from technological advances are not new.[143] Historically, advances in technology have had varied impacts on the labor market, with new technologies reducing demand for some skills and increasing demand for others.[144] The rapid advance of AI technologies and their application in multiple sectors of the economy have increased fears about possible job losses and spurred academic and government interest in studying potential impacts. Meanwhile, this has also led to concern that too few workers have AI expertise, both to work with AI in their jobs and to conduct AI R&D. Thus, discussions of AI and the U.S. workforce largely focus on two main issues: (1) the potential impact of AI and AI-driven automation on workers, including job displacement and job shifts; and (2) whether the United States has enough AI experts (people with advanced degrees in AI who work or teach in AI fields) for research, development, and application of AI across sectors, as well as teaching the next generation of AI experts.

Job Displacement and Skill Shifts

Economists and researchers are divided on possible answers to the question of how many jobs will be lost, gained, or changed, due partly or wholly to the development and application of AI technologies. Some analysts may argue that AI-related technologies are unprecedented in their speed of development, their range of applications, and the number of jobs they threaten, while others may argue that technology has a long history of displacing labor yet simultaneously creating new jobs, any net loss would be negligible, and the factors affecting the pace and extent of automation and AI adoption have not changed.[145] However, newly created jobs may be quite different from those eliminated and subsequently burden workers with the need to invest time, money, and relocation efforts in order to train for or acquire new jobs. A 2019 McKinsey Global Institute report that examined the impact of automation technologies on local economies and demographic groups stated, "While there could be positive net job growth at the national level, new jobs may not appear in the same places, and the occupational mix is changing. The challenge will be in addressing local mismatches and helping workers gain new skills."[146]

The potential impacts of AI technologies on the number and types of jobs that are or will be available are challenging to measure and predict for a variety of reasons.

---

[143] For a historical perspective, see for example, David H. Autor, "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives*, vol. 29, no. 3 (Summer 2015), pp. 3-30.

[144] Executive Office of the President (EOP), *Artificial Intelligence, Automation, and the Economy*, December 2016, p. 11, at https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF.

[145] For example, these perspectives are discussed in "Automation and Anxiety: Will Smarter Machines Cause Mass Unemployment," *The Economist*, June 23, 2016, at https://www.economist.com/special-report/2016/06/23/automation-and-anxiety.

[146] Susan Lund et al., *The Future of Work in America: People and Places, Today and Tomorrow*, McKinsey Global Institute, July 2019, available at https://www.mckinsey.com/featured-insights/future-of-work/the-future-of-work-in-america-people-and-places-today-and-tomorrow.

- First, definitions of AI and related technologies vary across industries, studies, and reports; further, potential job impacts from AI, computers, robots, and automation more generally are often conflated, making the specific workforce effects from AI technologies challenging to specify.
- Second, the numerous studies conducted to date vary in scope, including the labor sectors, populations, and countries assessed; the timeframes of predicted impacts; and the granularity of the datasets analyzed (e.g., whole occupations, specific tasks, or skillsets). One news article in 2018 attempted to compile all the available studies on how automation, AI, and robots could affect job losses or gains. The author summarized 19 studies that ranged in prediction dates (where specified) from 2016 to 2035, in jobs eliminated from 1.8 million to 1 billion, in jobs created from 1 million to 890 million, and in geographic focus from single countries (the United States or the United Kingdom) to worldwide. The author concluded that "there are about as many predictions as there are experts."[147] Further, many studies have relied on case studies and subjective assessments by experts.[148]
- Third, AI technologies are rapidly evolving, and it is difficult to predict what specific tasks they might be used to automate in the future, even in the short term. Some experts have asserted that "there is no widely shared agreement on the tasks where ML systems excel, and thus little agreement on the specific expected impacts on the workforce and on the economy more broadly."[149] And while AI is predicted to have greater displacement effects on higher skill professional and technical workers than earlier waves of automation, robust measures of current and future effects are still in development.[150]

While many reports and news stories related to job automation focus on worker displacement, some companies report using AI-enabled automation to perform jobs that are "dirty, dull, and dangerous," such as sorting at recycling facilities,[151] or to make up for labor shortages in the tight labor market. For example, some agriculture companies report developing autonomous systems to help make up for a shortage of farm workers.[152] Other companies making use of automation still report a high demand for employees. For example, Amazon reportedly expanded its workforce by 300,000 people since acquiring robotics company Kiva and deploying its robots in 2012 in its distribution centers. An employee overseeing robotics work at Amazon stated that "the biggest problem is not having enough people, and I don't think that is going to change."[153]

---

[147] Erin Winick, "Every Study We Could Find on What Automation Will Do to Jobs, in One Chart," *MIT Technology Review*, January 25, 2018, at https://www.technologyreview.com/s/610005/every-study-we-could-find-on-what-automation-will-do-to-jobs-in-one-chart/.

[148] Mark Muro, Jacob Whiton, and Robert Maxim, *What Jobs Are Affected by AI? Better-Paid, Better-Educated Workers Face the Most Exposure*, Brookings, November 2019, at https://www.brookings.edu/wp-content/uploads/2019/11/2019.11.20_BrookingsMetro_What-jobs-are-affected-by-AI_Report_Muro-Whiton-Maxim.pdf.

[149] Erik Brynjolfsson and Tom Mitchell, "What Can Machine Learning Do? Workforce Implications," *Science*, vol. 358, no. 6370 (2017), pp. 1530-1534.

[150] Michael Webb, "The Impact of Artificial Intelligence on the Labor Market," Stanford University Working Paper, July 2019.

[151] Bryn Nelson, "How Robots Are Reshaping One of the Dirtiest, Most Dangerous Jobs," *NBC News*, April 17, 2018, at https://www.nbcnews.com/mach/science/how-robots-are-reshaping-one-dirtiest-most-dangerous-jobs-ncna866771.

[152] Erin Winick, "New Autonomous Farm Wants to Produce Food Without Human Workers," *MIT Technology Review*, October 3, 2018, at https://www.technologyreview.com/s/612230/new-autonomous-farm-wants-to-produce-food-without-human-workers/.

[153] Cade Metz, "FedEx Follows Amazon into the Robotic Future," *New York Times*, March 18, 2018, at

While many studies over the past few years have discussed AI as part of automation technologies more broadly, some have begun trying to assess the AI- and ML-specific portions of potential impacts. Prior analyses looking more broadly at automation of job skills have generally found that lower-wage, blue-collar workers will be more affected. However, one 2018 study concluded that although most occupations have some tasks that could be automated using ML, there are few, if any, where all tasks are suitable for automation.[154] A 2019 study looking at AI-specific technologies found that (1) higher-wage, white-collar occupations and some agriculture and manufacturing positions may be the most exposed to AI disruptions; (2) AI seems likely to affect men, prime-age workers, and white and Asian American workers; and (3) large metropolitan areas with a concentration of high-tech industries and communities heavily involved in manufacturing are likely to experience the most AI-related disruption.[155] The authors caveat their work by noting that studies examining employment effects with any nuance are preliminary and that "the onset of AI will introduce new riddles into speculation about the future of work."[156] In general, recent studies indicate that most if not all occupations will be impacted by the introduction of AI and AI-enabled technologies in some way.

A 2020 report from the MIT Task Force on the Work of the Future asserted that the "momentous impacts of technological change are unfolding gradually," and that while applications and impacts from AI and robotics applications are coming, "they are not as close as some would fear."[157] The report discusses a variety of factors informing these findings, including that AI systems are still narrow and that policies, organizational cultures, economic incentives, and management practices can shape "the rate and manner in which firms develop and adopt technologies" beyond what is technologically possible.[158]

AI Expert Workforce

Tied to considerations of U.S. competitiveness, policymakers and stakeholders in academia and technology companies have expressed concerns about a lack of adequate AI expertise, not only for AI R&D and education in industry and academia, but also in the federal and congressional workforces. A September 2019 report highlighted several indicators of a tight market for AI talent, though the authors caveated their findings, noting that there is broad consensus in the field that talent shortages are substantial, but the exact extent is difficult to measure, and different organizations may publish very different estimates:[159]

https://www.nytimes.com/2018/03/18/technology/fedex-robots.html.

[154] Erik Brynjolfsson, Tom Mitchell, and Daniel Rock, "What Can Machines Learn and What Does It Mean for Occupations and the Economy?," *AEA Papers and Proceedings*, vol. 108 (May 2018), pp. 43-47, at https://pubs.aeaweb.org/doi/pdfplus/10.1257/pandp.20181019.

[155] Mark Muro, Jacob Whiton, and Robert Maxim, *What Jobs Are Affected by AI? Better-Paid, Better-Educated Workers Face the Most Exposure*, Brookings, November 2019, at https://www.brookings.edu/wp-content/uploads/2019/11/2019.11.20_BrookingsMetro_What-jobs-are-affected-by-AI_Report_Muro-Whiton-Maxim.pdf.

[156] Ibid, p. 22.

[157] David Autor, David Mindell, and Elisabeth Reynolds, *The Work of the Future: Building Better Jobs in an Age of Intelligent Machines*, Massachusetts Institute of Technology (MIT) Task Force on the Work of the Future, November 2020, pp. 5, 32-34, at https://workofthefuture.mit.edu/wp-content/uploads/2021/01/2020-Final-Report4.pdf.

[158] Ibid.

[159] Remco Zwetsloot, Roxanne Heston, and Zachary Arnold, *Strengthening the U.S. AI Workforce*, Center for Security and Emerging Technology, Georgetown University, September 2019, pp. 9-10. See the callout box, "What is the 'AI workforce,' and who counts as an 'AI expert'?", p. 3, for additional discussions of measuring the AI expert workforce.

- Job site statistics show that demand for workers far exceeds supply. For example, based on data from Burning Glass Technologies, job listings for AI skills have "grown significantly" from 2013 to 2020, with the total number of AI jobs posted in the United States above 300,000 in 2019 and 2020.[160] And as reported in April 2019, the market intelligence firm Element AI estimated that, in the United States, there were around 144,000 AI-related job openings and only about 26,000 developers and specialists seeking work.[161]
- The private sector is paying high salaries for workers with AI skills. For example, a 2018 news report stated that "even newly-minted Ph.D.s in machine learning and data science can make more than $300,000" at technology companies such as Google, Facebook, and Apple.[162]
- Subjective assessments from employers align with the indicators. For example, among firms surveyed by the World Economic Forum in 2020, most of which reported a desire to invest in AI, "skills gaps" and "inability to attract specialized talent" ranked among the top two barriers to the adoption of new technologies, especially when hiring for "emerging roles," including AI and ML specialists.[163]

Perhaps for this reason, some companies such as Google, Amazon, and Facebook, are recruiting professors while allowing them to retain positions at universities.[164] However, the details of these arrangements are important, as Oren Etzioni of the Allen Institute for Artificial Intelligence notes in an example from Facebook: "What are the ethics of a major corporation suddenly going after the entire [natural language processing] faculty in a computer science department? I believe their original offers had the faculty members spending 80 percent of their time at Facebook, which would not allow them time to carry out their educational responsibilities at [the University of Washington]." Some have referred to this as *eating the seed corn*, which could lead to less capacity to train future AI experts. Facebook disputed the claim, noting that while the relationship between academia and industry may be changing, the company is trying to be careful about not draining universities.[165] However, in a March 2019 survey of 111 AI researchers and university administrators by *Times Higher Education* and Microsoft, 89% said that it was "difficult" or "very difficult" to hire and retain AI experts.[166]

Other companies are collaborating with universities, such as Google's partnership with Princeton University to open an AI laboratory that will engage faculty members, graduate and undergraduate students, recent graduates, and software engineers. One of the collaborating faculty members, who previously split time between Princeton and Google, noted that it was an

---

[160] AI Index 2021, p. 86.

[161] As reported in Roberta Kwok, "Junior AI Researchers Are in Demand by Universities and Industry," *Nature*, April 23, 2019, at https://www.nature.com/articles/d41586-019-01248-w.

[162] Jeremy Kahn, "Sky-High Salaries Are the Weapons in the AI Talent War," *Bloomberg*, February 13, 2018, at https://www.bloomberg.com/news/articles/2018-02-13/in-the-war-for-ai-talent-sky-high-salaries-are-the-weapons.

[163] World Economic Forum, Center for the New Economy and Society, *The Future of Jobs Report 2020*, October 2020, pp. 27 and 35, at http://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf.

[164] Daniela Hernandez and Rachael King, "Universities' AI Talent Poached by Tech Giants," *Wall Street Journal*, November 24, 2016, at https://www.wsj.com/articles/universities-ai-talent-poached-by-tech-giants-1479999601.

[165] Alan Boyle, "FAIR Competition? Facebook Creates Official AI Labs in Seattle and Pittsburgh, Vying for Top Talent," *GeekWire*, May 5, 2018, at https://www.geekwire.com/2018/fair-competition-facebook-raises-status-ai-research-labs-seattle-pittsburgh/.

[166] As reported in Roberta Kwok, "Junior AI Researchers Are in Demand By Universities and Industry," *Nature*, April 23, 2019, at https://www.nature.com/articles/d41586-019-01248-w.

opportunity for those at Princeton to "benefit from exposure to real-world computing problems, and for Google to benefit from long-term, unconstrained academic research that Google may incorporate into future products."[167]

Within the federal government, the *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies* report asserted that "if we expect agencies to make responsible and smart use of AI, technical capacity must come from within" and "in-house expertise promotes AI tools that are better tailored to complex governance tasks and more likely to be designed and implemented in lawful, policy-compliant, and accountable ways."[168] To gain such expertise, the report states that "fully leveraging agency use of AI will require significant public investment to draw needed human capital."[169] Further, E.O. 13960 states that "agencies shall provide appropriate training to all agency personnel responsible for the design, development, acquisition, and use of AI." However, the March 2021 final report of the National Security Commission on Artificial Intelligence (NSCAI) states, "The human talent deficit is the government's most conspicuous AI deficit and the single greatest inhibitor to buying, building, and fielding AI-enabled technologies for national security purposes."[170]

Policy Considerations. Studies that attempt to identify the workforce effects of AI and ML technologies specifically, rather than those that address automation generally, conclude that there has been insufficient data collection and analyses specific to AI technologies and job skills conducted to fully understand the issue and inform policy decisions. For example, one study identified barriers that inhibit researchers from measuring the labor effects of AI, including (1) lack of high-quality data about the nature of work; (2) lack of empirically informed models of key microlevel processes (e.g., skill substitution and human-machine complementarity); and (3) insufficient understanding of how cognitive technologies interact with broader economic dynamics and institutional mechanisms.[171] The study asserted that overcoming such barriers requires improvements in the longitudinal and spatial resolution of data and refinements to data on workplace skills.[172] Another study, commissioned by the Bureau of Labor Statistics (BLS) to identify constructs that would complement BLS's existing products to assess the impact of automation on labor outcomes, echoed these findings. The BLS-commissioned study by Gallup states that "the primary lesson learned from [the] report is that researchers and, by extension, policymakers lack the data necessary to fully understand how new technologies impact the labor market" and identified gaps in BLS data products, specifically with regards to the classification of skills, task performance, and the adoption of new technologies.[173]

[167] Steven Schultz, "Google to Open Artificial Intelligence Lab in Princeton and Collaborate with University Researchers," Princeton University news communication, December 18, 2018, at https://www.princeton.edu/news/2018/12/18/google-open-artificial-intelligence-lab-princeton-and-collaborate-university.

[168] ACUS report 2020, p. 7.

[169] Ibid.

[170] National Security Commission on Artificial Intelligence, *Final Report*, March 2021, p. 3, at https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf (hereinafter, "NSCAI 2021 Final Report").

[171] Morgan R. Frank et al., "Toward Understanding the Impact of Artificial Intelligence on Labor," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 14 (April 2, 2019), pp. 6531-6539.

[172] Ibid.

[173] Jenny Marler, Gallup Project Director, Assessing the Impact of New Technologies on the Labor Market: Key Constructs, Gaps, and Data Collection Strategies for the Bureau of Labor Statistics, Contract No: GS-00F-0078M**,** February 7, 2020, pp. 3, 25 (hereinafter referred to as the Gallup study), at https://www.bls.gov/bls/congressional-reports/assessing-the-impact-of-new-technologies-on-the-labor-market.pdf.

Some experts emphasize training people for skills and jobs that will be in high demand even with implementation of AI technologies, such as skills needed in management and personal interactions, two areas for which AI is not well suited.[174] Stakeholders have also asserted that a focus on lifelong learning and programs to retrain and upskill workers will be important for addressing skill shifts related to deployment of AI technologies.[175] In one 2017 survey of 300 C-suite and senior executives about their AI strategies, 82% of leaders planned to implement AI in the next three years, but only 38% provided programs aimed at reskilling employees to work with the technology.[176] Still other experts assert that "the concern should not be about the number of jobs, but whether those are jobs that can support a reasonable standard of living and what set of people have access to them."[177]

In response to these issues, some policy questions and considerations for Congress may include the following:

- What types of granular labor data are needed to better inform analyses and identify key skills for future jobs, and how might the federal government help gather and disseminate such information?
- In conjunction with efforts by employers and educators, what is the appropriate role of the federal government in supporting the reskilling or upskilling of employees for whom certain tasks or their entire jobs will be shifted or displaced? Are federal programs to assist workers sufficient to help address potential workforce shifts? How can federal direction of workforce support programs balance providing AI-specific legislative direction while allowing states and localities flexibility to meet their specific workforce needs?
- For those federal offices and agencies facing a shortage of technical expertise in AI, what are the best options to attract and retain talent? For example, former Secretary of Defense Robert Work has argued for the development of an AI training corps—similar to the CyberCorps program[178] (educational training in exchange for expert work for the federal government, but where workers could keep their regular jobs).[179]

---

[174] David Rotman, "Obama Economist: We're Not Preparing Workers for Changing Jobs," *MIT Technology Review*, June 4, 2018, at https://www.technologyreview.com/s/611297/obama-economist-were-not-preparing-workers-for-changing-jobs/; video of Jason Furman's talk at the 2018 EmTech conference, covered in the article, can be found at https://events.technologyreview.com/video/watch/jason-furman-harvard-automation-future-work/.

[175] James Manyika et al., *A Future That Works: Automation, Employment, and Productivity*, McKinsey Global Institute, January 2017, available at https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works; and Joseph E. Aoun, *Robot-Proof: Higher Education in the Age of Artificial Intelligence* (Cambridge, MA: MIT Press, 2017). Generally, reskilling refers to learning new skills for a different job or occupation, while upskilling refers to learning new skills for growth within an existing job or occupation.

[176] Genpact, "Is Your Business AI-Ready?," 2017, at http://www.genpact.com/downloadable-content/insight/is-your-business-ai-ready.pdf.

[177] David Autor, "No, Robots Won't Take All the Jobs," *Brookings Creative Lab*, March 12, 2018, at https://www.youtube.com/watch?v=SrprBJf7Nd4 (video discussion of the paper, David Autor and Anna Salomons, "Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share," Brookings Papers on Economic Activity, Spring 2018, at https://www.brookings.edu/bpea-articles/is-automation-labor-displacing-productivity-growth-employment-and-the-labor-share/).

[178] See for example, the CyberCorps Scholarship for Service program at https://www.sfs.opm.gov/.

[179] David Ignatius, "China's Application of AI Should Be a Sputnik Moment for the U.S. but Will It Be?," *Washington Post*, November 6, 2018, at https://www.washingtonpost.com/opinions/chinas-application-of-ai-should-be-a-sputnik-moment-for-the-us-but-will-it-be/2018/11/06/69132de4-e204-11e8-b759-3d88a5ce9e19_story.html.

- In addition to developing internal expertise, how might federal agencies and executive offices expand access to outside expertise, as from academia, industry, and nonprofit groups? For example, the NSCAI 2021 Final Report recommends establishing a civilian National Reserve Digital Corps modeled after the military reserve's commitment and incentive structure.[180]

In order to address the dearth of data on the potential impacts of AI on the workforce, Congress may consider various actions. The FY2021 NDAA calls for the commissioning of a study by the National Academies of Sciences, Engineering, and Medicine on the current and future impact of AI on the workforce of the United States across sectors, including addressing research gaps and data needed to better understand workforce impacts. The study may yield useful information to inform the debate and future policy options; the final report is due more than two years from enactment, which occurred in January 2021. During that time, Congress may hold hearings to obtain related information on new or updated data collection and research at federal agencies in response to prior studies. Further, Congress may direct federal agencies to begin collecting additional information to fill data gaps identified in prior research, such as in the Gallup study for BLS.

Should Congress decide to assist federal agencies in attracting outside expertise and developing internal expertise in AI, a variety of policy responses have been discussed by stakeholders. For example, Congress may consider directing federal agencies to develop or expand on scholarship-for-service (SFS) programs to attract new AI talent to federal service. However, simply expanding the number of offerings may not result in more students participating—such programs have been criticized for being difficult to find online, being spread across multiple and possibly outdated agency websites, and not supporting continued professional development once a student is employed in the federal government.[181] While SFS programs have had reportedly high placement rates for graduates—94% for CyberCorps graduates in 2016—some critics have expressed discomfort with the repayment requirements for students who enter the program but leave before completing their degree or federal service requirement.[182] Further challenges for growing a federal workforce in AI include higher salaries for comparable jobs in the private sector and time-consuming and opaque hiring practices. Thus, Congress may consider directing agencies to take actions to improve the recruitment and retention of AI experts, including through the establishment or modification of federal programs such as SFS.

Developing internal expertise at agencies to not only develop, but use, understandable and transparent AI systems may have multiple benefits for agencies. For example, agency experts likely have a deeper, more nuanced understanding of the technical needs and challenges at their agency for which an AI system is developed or tailored. Further, by developing their own AI systems, agencies may be better able to create understandable, transparent, and accountable systems, in contrast to the estimated 33% of federal AI systems that are built by external contractors using proprietary software and obtained through the federal procurement process.[183] Congress may consider ways to support or augment AI expertise within the existing federal

[180] NSCAI 2021 Final Report, p. 125.

[181] Cindy Martinez, "Saving the Federal Cyber and AI Workforce from Obsolescence: How to Attract and Retain a New Generation," *FedScoop*, December 22, 2020, at https://www.fedscoop.com/saving-federal-cyber-ai-workforce-obsolescence-attract-retain-new-generation/.

[182] For additional discussions of SFS programs and the federal workforce in the context of cybersecurity, see CRS In Focus IF10654, *Challenges in Cybersecurity Education and Workforce Development*, by Boris Granovskiy.

[183] Administrative Conference of the United States, Office of the Chairman Projects, "Artificial Intelligence in Federal Agencies," February 2020, p. 88-98.

workforce through the establishment of federal advisory committees and directing agencies to develop internal training programs.

International Competition and Federal Investment in AI R&D

According to the National Science Board's *Science and Engineering Indicators for 2020*, the United States and China lead in research and commercialization of AI technologies, though business adoption of AI is occurring across the world.[184] Numerous international governments have initiated activities focused on AI (e.g., task forces, research activities, discussion papers), and dozens have released national AI strategies, though these vary in scope.[185] Further, multiple countries are cooperating in international AI initiatives. For example, the United States and other Organisation for Economic Co-operation and Development (OECD) member countries committed to common AI principles in May 2019.[186] Building on the commitment to these principles, the United States and 14 other countries launched the Global Partnership on AI in June 2020 to bring together expertise from a range of stakeholders "with the goal of bridging the gap between the theory and practice of AI."[187] In September 2020, the United States and the United Kingdom signed a declaration of cooperation in AI R&D.[188]

Public investments in AI R&D vary widely by country. In the United States, as previously noted, FY2020 funding for AI activities at defense and non-defense agencies was approximately $4 billion and $1.1 billion, respectively. In comparison, a recent report from the Center for Security and Emerging Technology at Georgetown University estimated that Chinese government spending on AI R&D in 2018 was on the order of a few billion dollars.[189] Though a substantial amount, this is less than the estimate of tens of billions that others have suggested. The European Union previously communicated a commitment to increase investments from $500 million to $1.5 billion by the end of 2020. In 2018, Germany and France pledged €3 billion and €1.5 billion, respectively, for AI investments by the end of 2020, and Canada previously committed to spending $125 million over five years.[190]

[184] National Science Board, National Science Foundation, "Production and Trade of Knowledge- and Technology-Intensive Industries," *Science and Engineering Indicators 2020*, NSB-2020-5, p. 55, at https://ncses.nsf.gov/pubs/nsb20205/.

[185] One of the most comprehensive efforts to compile information on AI initiatives across countries has been conducted through the Organisation for Economic Co-operation and Development's (OECD's) AI Policy Observatory, at https://oecd.ai/.

[186] OECD, *Recommendation of the Council on Artificial Intelligence*, adopted on May 21, 2019, at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[187] National Artificial Intelligence Initiative Office, "Global Partnership on AI," at https://www.ai.gov/strategic-pillars/international-cooperation/#Global-Partnership-on-AI. For additional information about the Global Partnership on Artificial Intelligence, see https://gpai.ai/.

[188] U.S. Department of State, "Declaration of the United States of America and the United Kingdom of Great Britain and Northern Ireland on Cooperation in Artificial Intelligence Research and Development: A Shared Vision for Driving Technological Breakthroughs in Artificial Intelligence," September 25, 2020, at https://www.state.gov/declaration-of-the-united-states-of-america-and-the-united-kingdom-of-great-britain-and-northern-ireland-on-cooperation-in-artificial-intelligence-research-and-development-a-shared-vision-for-driving/.

[189] Ashwin Acharya and Zachary Arnold, *Chinese Public AI R&D Spending: Provisional Findings*, Center for Security and Emerging Technology, Georgetown University, December 2019.

[190] Information on the status of these investments is unknown. As previously noted, it is important to keep in mind that reliable cross-country measures on public investments are difficult to obtain for a variety of reasons, including varying levels of reporting, and the range of measurements that countries could use to tally spending.

Policy Considerations. The appropriate level for U.S. federal R&D support, the nature of the R&D investments, such as basic versus applied research, as well as the most effective additional mechanisms to support innovation, such as prize competition incentives and public-private partnerships, remain areas of discussion among lawmakers.

Historical considerations of international competition in science and technology have led to prior recommendations for increased federal funding of research, particularly in the physical sciences and engineering (PS&E).[191] For example, the America COMPETES Act (P.L. 110-69) in 2007 and the America COMPETES Reauthorization Act of 2010 (P.L. 111-358) were originally enacted to address concerns that the United States could lose its advantage in scientific and technological innovation. The COMPETES Acts included authorizations of appropriations in line with doubling research in PS&E, including doubling NSF's budget. Appropriations for the COMPETES Acts activities never reached authorized levels, and opposition to the efforts included various perspectives, including a preference for alternative federal approaches to support innovation, such as research tax credits or reducing regulatory costs, as well as a concern about the national debt.[192]

More recently, regarding federal funding and support for AI R&D, some stakeholders assert that the federal government should invest more money and direct structural or programmatic changes to certain R&D agencies to promote U.S. technological primacy, particularly in key areas of emerging technologies such as AI. For example, the President's Council of Advisors on Science and Technology (PCAST) released recommendations in June 2020 on strengthening U.S. American Leadership in industries of the future, which included growing federal investment in AI R&D by a factor of 10 over 10 years (e.g., increase non-defense R&D from $1 billion in FY2020 to $10 billion in FY2030).[193]

The National AI Initiative Act, passed in the FY2021 NDAA, authorized appropriations for AI activities at NSF, NIST, and DOE for FY2021-FY2025. In the 117th Congress, the Endless Frontier Act (S. 1260) would redesignate the NSF as the National Science and Technology Foundation, establishing a Directorate for Technology and authorizing appropriations of $100 billion over five years for the new directorate.[194] The final report of the National Security Commission on AI recommends scaling and coordinating federal AI R&D funding, including through establishing a National Technology Foundation as a sister agency to the NSF "to provide the means to move science more aggressively into engineering and scale innovative ideas into reality"; funding AI R&D at compounding levels; and establishing additional National AI Research Institutes.[195] Congress considers the appropriations for these authorities as part of its annual discretionary appropriations process and enacted amounts may or may not match the authorized levels.

---

[191] National Academies of Sciences, Engineering, and Medicine, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, 2007, at https://doi.org/10.17226/11463.

[192] For additional discussions of the America COMPETES Acts and efforts to double federal PS&E funding, see CRS Report R41951, *An Analysis of Efforts to Double Federal Funding for Physical Sciences and Engineering Research*, by John F. Sargent Jr.

[193] President's Council of Advisors on Science and Technology (PCAST), *Recommendations for Strengthening American Leadership in Industries of the Future*, June 2020, p. 6, at https://science.osti.gov/-/media/_/pdf/about/pcast/202006/PCAST_June_2020_Report.pdf?la=en&hash=019A4F17C79FDEE5005C51D3D6CAC81FB31E3ABC.

[194] For comparison, FY2021 appropriations for NSF were approximately $8.5 billion total. The Endless Frontier Act was first introduced in the 116th Congress (S. 3832 and H.R. 6978).

[195] NSCAI 2021 Final Report, p. 435.

An additional consideration, given the R&D engagement in the private sector, is the extent to which the federal government might leverage private funding through expanding public-private partnerships. In the 2019 update to the *National AI R&D Strategic Plan*, expanding public-private partnerships to accelerate advances in AI was a new, additional strategy.

Standards Development

AI standards development became an area of increasing interest for the Trump Administration and the 116th Congress, for both domestic R&D and international competitiveness reasons. The 2019 *National AI R&D Strategic Plan* noted that "development and adoption of best practices and standards in documenting dataset and model provenance will enhance trustworthiness and responsible use of AI technologies."[196] E.O. 13859 aimed to "Ensure that technical standards … reflect Federal priorities for innovation, public trust, and public confidence in systems that use AI technologies … and develop international standards to promote and protect those priorities." In response, NIST produced the *Plan for Federal Engagement in Developing Technical Standards and Related Tools* (AI Standards Plan) in August 2019. The plan identifies nine areas of focus for AI standards: concepts and terminology; data and knowledge; human interactions; metrics; networking; performance testing and reporting methodology; safety; risk management; and trustworthiness.[197]

The standards development process in the United States is predominantly a voluntary, consensus-based effort, driven by the private sector, including through Standards Development Organizations (SDOs). NIST (with other federal agencies, as appropriate) is a participant and facilitator, providing agency requirements to standards projects and technical expertise to standards development, incorporating voluntary standards into policies and regulations, and citing standards in agency procurements.[198] Standards can be horizontal (i.e., used across many applications and industries), or vertical (i.e., developed for specific application areas such as healthcare or transportation). Further, nontechnical standards can be important to inform policy and human decisionmaking (e.g., standards for governance and privacy), and "standards should be complemented by an array of related tools," such as standardized datasets with metadata; benchmarks; testing methodologies; metrics; testbeds; and tools for accountability and auditing.[199] The AI Standards Plan notes that "While there is broad agreement that [federal policies and principles, including those that address societal and ethical issues, governance, and privacy] must factor into AI standards, it is not clear how that should be done and whether there is yet sufficient scientific and technical basis to develop those standards provisions."[200]

[196] NSTC Select Committee on Artificial Intelligence 2019 AI R&D Strategic Plan, p. 28.

[197] National Institute of Standards and Technology, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, August 9, 2019, pp. 3, 10-12. The plan further states that "Trustworthiness standards include guidance and requirements for accuracy, explainability, resiliency, safety, reliability, objectivity, and security."

[198] Ibid, "How Are Technical Standards Developed?" p. 9. The document also includes a list of SDOs that are developing AI standards in Appendix II. For additional information about NIST, including certain statutory authorities, see CRS Report R43908, *The National Institute of Standards and Technology: An Appropriations Overview*, by John F. Sargent Jr.

[199] Ibid, pp. 13-14.

[200] Ibid, p. 4. Social and ethical issues are discussed in the following section, "Ethics, Bias, Fairness, and Transparency."

Standards development is not only a national but an international effort, involving the work of such entities as the International Organization for Standardization (ISO).[201] The U.S. government and other stakeholders have expressed concern about China's attempts to lead the international AI standards development efforts. China has already laid out some of these plans in white papers and is expected to release a 15-year plan to set global standards for next-generation technologies, including AI, as part of its "China Standards 2035" plan.[202] Concerns about China's focus on standards setting, particularly if the United States does not lead in these efforts, include the following.

- **Potential economic losses.** The NIST AI Standards Plan highlights this concern, stating, "AI standards developed without the appropriate level and type of involvement may exclude or disadvantage U.S.-based companies in the marketplace as well as U.S. government agencies."[203]
- **Threats to democratic norms and values.** Members of the National Security Commission on AI have expressed concern that "AI is being used in ways that are antithetical to American values. In China, AI is used as a tool for centralizing power at the expense of individual rights. The Chinese government is amassing the personal data of its people, using facial recognition software to stifle dissent and repress minorities, and exporting its surveillance technology abroad."[204] The ability of those countries leading in international standards setting to impart their societal and cultural values, such as data privacy and respect for civil liberties, into the process and outcomes, has led to concerns about China's successes in increasing its leadership positions in international standards-making bodies.[205] As NIST has stated, "standards flow from principles, and a first step toward standardization will be reaching broad consensus on a core set of AI principles."[206]

These points are discussed in greater detail in the U.S.-China Economic and Security Review Commission's 2020 annual report to Congress, which states

> In contrast to the United States, where technical standards are developed by industry in response to commercial need and adopted by consensus, Chinese state agencies formulate standards and use them to advance industrial and foreign policy objectives. Historically, Beijing has prioritized developing mandatory and unique domestic technical standards as a barrier to foreign firms' market entry and to help grow domestic industry. Now, it is also coordinating industrial policy and diplomatic strategy to expand its influence in

[201] Information about the ISO Committee on AI can be found at https://www.iso.org/committee/6794475.html.

[202] The Center for Security and Emerging Technology (CSET) at Georgetown University has provided a translation of China's *Artificial Intelligence Security Standardization White Paper*, 2019, at https://cset.georgetown.edu/wp-content/uploads/t0121_AI_security_standardization_white_paper_EN.pdf; regarding the forthcoming "China Standards 2035," see Arjun Kharpal, "Power Is 'Up for Grabs': Behind China's Plan to Shape the Future of Next-Generation Tech," *CNBC*, April 26, 2020, at https://www.cnbc.com/2020/04/27/china-standards-2035-explained.html.

[203] NIST, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, p. 19.

[204] Eric Schmidt and Bob Work, "The US Is in Danger of Losing Its Global Leadership in AI," *The Hill*, December 5, 2019, at https://thehill.com/blogs/congress-blog/technology/473273-the-us-is-in-danger-of-losing-its-global-leadership-in-ai.

[205] U.S.-China Economic and Security Review Commission, *2020 Annual Report to Congress*, December 2020, p. 107, at https://www.uscc.gov/files/001592.

[206] NIST, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, p. 15.

international standards-making bodies, both to increase adoption of Chinese technology abroad and to influence norms for how technology is applied.[207]

Policy Considerations. Such concerns have generated various recommendations for robust domestic and international standards setting efforts. The AI Standards Plan included numerous recommendations to support U.S. leadership in AI standards development:

- Bolster AI standards-related knowledge, leadership, and coordination among federal agencies, including by:
    - designating a Standards Coordinator within the NSTC's MLAI Subcommittee, and
    - developing clear career development and promotion paths that encourage participation and expertise in AI standards and development.
- Promote focused research to advance and accelerate broader exploration and understanding of how aspects of trustworthiness can be practically incorporated within standards and standards-related tools, including through supporting research to develop metrics, data sets, and risk management strategies for AI.
- Support and expand public-private partnerships to develop and use AI standards and related tools to advance reliable, robust, and trustworthy AI.
- Strategically engage with international parties to advance AI standards for U.S. economic and national security needs, including through accelerating information exchange with "like minded countries" through international partnerships.[208]

In the FY2021 NDAA (Section 5301), Congress established as a mission that NIST advance collaborative frameworks, standards, guidelines; authorized NIST to work on associated methods and techniques for AI; and directed that NIST support the development of a risk-management framework for trustworthy AI systems. NIST is further directed to develop guidance and best practices for data set documentation and data sharing among industry, federally funded research and development centers, and federal agencies, including options for partnerships with universities and nonprofits. Congress may consider oversight activities to monitor the implementation of these provisions and provide subsequent direction to NIST and other federal agencies.

Ethics, Bias, Fairness, and Transparency

Along with interest in technical advances, researchers, companies, and policymakers are expressing growing concern and interest in what has been called the ethical evolution of AI, including questions about bias, fairness, and algorithm transparency. Broadly, who defines ethics and who enforces ethical design and use?[209] What constitutes an ethical decision may vary by individual, culture, economics, and geography.[210] As some analysts have asserted, "AI is only as

[207] U.S.-China Economic and Security Review Commission, *2020 Annual Report to Congress*, December 2020, p. 106, at https://www.uscc.gov/files/001592.

[208] NIST, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, pp. 4-6.

[209] Karen Hao, "Establishing an AI Code of Ethics Will Be Harder Than People Think," *MIT Technology Review*, October 21, 2018, at https://www.technologyreview.com/2018/10/21/139647/establishing-an-ai-code-of-ethics-will-be-harder-than-people-think/.

[210] Edmond Awad et al., "The Moral Machine Experiment," *Nature*, October 24, 2018.

good as the information and values of the programmers who design it, and their biases can ultimately lead to both flaws in the technology and amplified biases in the real world."[211]

Just as there are many ways of considering what is ethical in AI, "researchers studying bias in algorithms say there are many ways of defining fairness, which are sometimes contradictory," having inherent tradeoffs.[212] (For example, one computer scientist presented at the 2018 Fairness, Accountability, and Transparency (FAT*) Conference on "21 fairness definitions and their politics."[213]) The box below presents an example of the challenges of defining fairness in the criminal justice system. For some, such cases highlight the need for agencies to improve their internal processes for assessing algorithmic tools and develop training for their staff to be able not only to evaluate such tools, but also to provide developers with publicly-available metrics for fairness.[214]

Sector Example: Defining Fairness in Criminal Justice

In 2016, a team at ProPublica investigated proprietary software called COMPAS that is used during sentencing to assign defendants in the criminal justice system with risk scores, from 1 to 10, for committing another crime within two years if released (i.e., the likelihood of recidivism). The ProPublica team claimed that the algorithm was biased, because there were a disproportionate number of false positives for black defendants—people identified as high risk who were not subsequently charged with another crime (one measure of an "error rate").[215] The developers countered that the algorithm was not biased, because it was equally good at predicting whether a white or a black defendant classified as high risk would reoffend, a measure called "predictive parity." In other words, ProPublica and the developers of COMPAS were using different measures to try to conclude whether the software was fair.

Subsequent research into these analyses found that not all criteria for fairness can be satisfied when recidivism prevalence differs across groups and that disparate impact—which the researcher defined as referring "to settings where a penalty policy has unintended disproportionate adverse impact on a particular group"—may result even if a prediction instrument is fair with respect to certain criteria.[216] The researcher—citing a large body of literature showing that data-driven risk assessment instruments tend to be more accurate than professional human judgements—concluded that data-driven approaches should not be abandoned but rather proven to be free of the kinds of biases that could lead to disparate impacts in the specific contexts in which they are applied.[217]

For a more in-depth discussion of this topic, see "Concerns About Bias in Risk and Needs Assessments" in CRS Report R44087, *Risk and Needs Assessment in the Federal Prison System*, by Nathan James.

The U.S. *National AI R&D Strategic Plan* also discusses the challenges and potential approaches to designing and building ethical AI. The plan echoes concerns about the susceptibility of data-

[211] Andre M. Perry and Nicol Turner Lee, "AI Is Coming to Schools, and If We're Not Careful, So Will Its Biases," *Brookings*, September 26, 2019, at https://www.brookings.edu/blog/the-avenue/2019/09/26/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases.

[212] Rachel Courtland, "Bias Detectives: The Researchers Striving to Make Algorithms Fair," *Nature News Feature*, vol. 558 (June 20, 2018), pp. 357-360 (hereinafter, "Courtland, 2018").

[213] Arvind Narayanan, "Translation Tutorial: 21 Fairness Definitions and Their Politics," Fairness, Accountability, and Transparency (FAT*) Conference, February 23, 2018; abstract and video available at https://www.youtube.com/watch?v=jIXIuYdnyyk. As noted on the conference website (https://facctconference.org/2018/program.html), "In 2018, the conference's name was FAT* and the proceedings were published in the Journal of Machine Learning Research. The conference affiliated with ACM in 2019, and changed its name to ACM FAccT immediately following the 2020 conference."

[214] Courtland, 2018.

[215] Julia Angwin et al., "Machine Bias," *ProPublica*, May 23, 2016, at https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[216] Alexandra Chouldechova, "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments," *Big Data*, vol. 5, no. 2 (June 2017), pp. 153-163, at https://pubmed.ncbi.nlm.nih.gov/28632438/.

[217] Ibid.

intensive AI algorithms to error and misuse without the proper collection and use of data to train the systems. It calls for researchers to design systems so that their actions and decisionmaking are more transparent and easily interpretable, and they can be examined for bias. The plan further states, "Ethics is inherently a philosophical question while AI technology depends on, and is limited by, engineering.… However, acceptable ethics reference frameworks can be developed to guide AI system reasoning and decisionmaking in order to explain and justify its conclusions and actions." To achieve these goals, the plan notes that there is a need for multidisciplinary, fundamental research in designing architectures for AI systems to incorporate ethical reasoning.[218]

While such fundamental research is being conducted, and while various groups work on developing standards and benchmarks for evaluating algorithms, some stakeholders have called for a risk-based, sector-specific approach to considering uses and potential regulations for AI algorithms. For example, some have called for more initial and ongoing testing and evaluation of algorithms and AI technologies for potential bias that directly impact U.S. citizens' lives and livelihoods (e.g., through healthcare or hiring systems)—sometimes referred to as "high risk" or "systems critical" uses.[219] Some Members of Congress have previously requested information from federal agencies about their use of AI, such as the use of facial recognition technology in law enforcement, and how the agencies balance the potential to solve crimes and catch criminals with the potential risks to privacy and civil rights.[220]

Types of Bias

Definitions and understanding of terms such as bias and fairness can vary by discipline (e.g. technologists vs. lawyers vs. civil society), type (e.g., statistical vs. social bias), and scope (e.g. individual vs. systemic/structural). Further, there are various types of bias, and bias can show up in algorithms, including AI algorithms, in a variety of ways, including in the data, within the system, and from the people designing and using the system.

There is significant concern that biases and errors in datasets used to train AI systems will result in outcomes that reflect, and possibly amplify, those biases. For example, using a dataset that has historical inequities engrained in it—such as past employment or access to credit, both of which have a history of racial discrimination—can perpetuate bias and inequity. Limited datasets that are not representative of the population to which they will be applied may lack generalizability and subsequently not work equally well for everyone. For example, some facial analysis software has been shown to have significant gender and skin color classification bias, often accurately identifying white males while failing to accurately classify darker female faces one in three times.[221] Another study found that two prominent research-image collections display gender bias in their depiction of activities such as cooking and sports; ML algorithms trained on these collections not only mirrored, but amplified, these biases.[222]

[218] NSTC Select Committee on Artificial Intelligence 2019 AI R&D Strategic Plan, pp. 21-22.

[219] See, for example, a discussion of racial bias in health care decisionmaking software used by hospitals in Heidi Leford, "Millions of Black People Affected by Racial Bias in Health-Care Algorithms," *Nature News*, vol. 574 (October 26, 2019), pp. 608-609.

[220] See for example, Letter from Senator Ron Wyden et al. to Gene L. Dorado, Comptroller General of the United States, July 31, 2018, at https://www.wyden.senate.gov/download/07312018-gao-facial-recognition-request.

[221] See work conducted by the Gender Shades project by Joy Buolamwini at the Massachusetts Institute of Technology's (MIT's) Media Lab, at https://www.media.mit.edu/projects/gender-shades/overview/.

[222] Jieyu Zhao et al., "Men Also Like Shopping: Reducing Gender Bias Amplification Using Corpus-level Constraints," *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark,

Certain variables may reflect societal inequities and stand in as proxies for protected classes of data (e.g., race, sex), inadvertently perpetuating prohibited discriminatory practices. Practices that result in disparate impacts may violate various laws, such as equal credit or employment opportunity laws.[223] For example, the algorithm in the COMPAS tool (see the criminal justice box above) purports to predict the risk of future criminal activity, but it relies on inputs such as arrest history; variations in historical policing could reflect over-policing of certain communities leading to a higher number of arrests and higher correlation with crime while not accurately reflecting the likelihood of recidivism.[224]

While the above example is for a relatively simple statistical algorithm, the "black box" problem with many complex AI systems may make assessments of such bias harder to evaluate and correct. Identifying and addressing machine bias is a challenging problem, fueling a growing subfield of AI research. In trying to address pronoun gender bias in its "smart compose" feature, which automatically completes sentences for users as they type, Google opted to ban the use of gendered pronouns, stating that currently, "the only reliable technique we have is to be conservative."[225]

Beyond these arguably unintentional instances of bias perpetuation and amplification, concerns have been raised about the potential for intentional introduction of bias into algorithms through the release or use of manipulated training data.[226]

Additionally, what has been termed *automation bias* can occur when people trust the interpretations of an automated system over their own senses and instincts, expecting the algorithmic outcomes to be objective calculations since they are being performed by a computer, rather than an individual person making a decisions.[227] However, even some particularly complex AI algorithms such as deep neural networks that can work exceedingly well the majority of the time can have catastrophic failures, breaking in unpredictable ways.[228] For example, researchers have demonstrated that placing black and white stickers on a stop sign can cause a neural network to misclassify the sign—for example, as a 45 miles-per-hour speed limit sign—over 80% of the time.[229]

Broadly, the debate around how to address bias and ethics in decisionmaking algorithms has resulted in calls for additional transparency, which raises its own sets of opportunities and

---

September 7, 2017, pp. 2979-2989, at https://www.aclweb.org/anthology/D17-1323.

[223] Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion?*, January 2016, p. 19; "While specific disparate impact standards vary depending on the applicable law, in general, disparate impact occurs when a company employs facially neutral policies or practices that have a disproportionate adverse effect or impact on a protected class."

[224] Courtland, 2018.

[225] Paresh Dave, "Fearful of Bias, Google Blocks Gender-Based Pronouns from New AI Tool," *Reuters*, November 27, 2018, at https://www.reuters.com/article/us-alphabet-google-ai-gender/fearful-of-bias-google-blocks-gender-based-pronouns-from-new-ai-tool-idUSKCN1NW0EF. The article further notes that gender-based pronoun biases are a widespread challenge for companies using AI for features such as natural language generation (NLG) and translation services.

[226] Douglas Yeung, "When AI Misjudgment Is Not an Accident," *Scientific American*, October 19, 2018, at https://blogs.scientificamerican.com/observations/when-ai-misjudgment-is-not-an-accident.

[227] John Zerilli et al., "Algorithmic Decision-Making and the Control Problem," *Minds and Machines*, vol. 29 (2019), pp. 555-578, at https://link.springer.com/article/10.1007/s11023-019-09513-7.

[228] Douglas Heaven, "Why Deep-Learning AIs Are So Easy to Fool," *Nature*, vol. 574 (October 9, 2019), pp. 163-166, at https://www.nature.com/articles/d41586-019-03013-5.

[229] Kevin Eykholt et al., "Robust Physical-World Attacks on Deep Learning Visual Classification," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, June 18-23, 2018, at https://ieeexplore.ieee.org/document/8578273.

challenges and questions about how best to enhance transparency. On one hand, engaging a broader set of stakeholders and providing information to those affected and journalists investigating the tools generally helps to foster trust and lead to fewer problems with bias and inequities. However, just providing all of the parameters of a model may not lead to better information about how it works. Further, providing too much information may allow people to game the system, and could provide a disincentive for private sector developers wishing to license their software. One compromise that has been proposed in this situation is to require confidential third-party auditing of proprietary software with publicly released results of such audits.[230]

Policy Considerations. Some considerations for potential policy responses to these issues include:

- Whether and how to increase access to public datasets to train AI systems for use in the public and private sectors;
- Requirements for auditing and/or disclosing AI algorithms—particularly in high-impact areas such as social services, criminal justice, and healthcare—and direction to NIST to facilitate related standards and certifications for third-party auditors;
- Mechanisms for recourse when people are subject to decisions in high-impact areas in which AI systems were used;
- Facilitating the growth of multidisciplinary and diverse teams of experts for developing and training AI systems, including having people who will be using and affected by the systems as part of the design conversations;
- Encouraging training for AI researchers and designers in thinking about and designing systems that improve fairness, transparency, and accountability; and
- Whether to continue or expand investments into AI R&D broadly and for more narrowly specified areas, such as those that facilitate transparency and auditability (e.g., explainable AI).

Author Information

Laurie A. Harris
Analyst in Science and Technology Policy

[230] See, for example, Oren Etzioni, and Michael Li, "High-Stakes AI Decisions Need to Be Automatically Audited," *Wired*, July 18, 2019, at https://www.wired.com/story/ai-needs-to-be-audited/.

Disclaimer