Artificial Intelligence:
A Policy-Oriented Introduction

Wilson Center

By Anne Bowser, Michael Sloan, Pietro Michelucci and Eleonore Pauwels

Technological advances in artificial intelligence (AI) promise to be pervasive, with impacts and ramifications in health, economics, security and governance. In combination with other emerging and converging technologies, AI has the potential to transform our society through better decision-making and improvements to the human condition. But, without adequate risk assessment and mitigation, AI may pose a threat to existing vulnerabilities in our defenses, economic systems, and social structures. Recognizing the increasing integration of technology in society, this policy brief grounds the present excitement around AI in an objective analysis of capability trends before summarizing perceived benefits and risks. It also introduces an emerging sub-field of AI known as Human Computation, which can help achieve future AI capabilities by strategically inserting humans in the loop where pure AI still falls short. Policy recommendations suggest how to maximize the benefits and minimize the risks for science and society, particularly by incorporating human participation into complex socio-technical systems to ensure the safe and equitable development of automated intelligence.

THE RESEARCH, DEVELOPMENT, AND POLICY CONTEXT

Artificial Intelligence (AI) is an academic discipline that began approximately sixty years ago. At the time, researchers believed that if a computer could be programed to solve basic mathematical problems it should be feasible, with enough information, to have machines exhibit a broader humanlike intelligence.

The term "Artificial Intelligence" was coined by the computer scientist John McCarthy in 1955.[1] In 1956, McCarthy organized the now famous Dartmouth Conference – formally the Dartmouth Summer Research Project on Artificial Intelligence – now widely recognized as the origin for the discipline. This conference helped initiate a critical period in AI research, from the 1950s to the 1970s. Early developments included heuristic search and machine learning with the aim of achieving broader objectives such as computer vision, natural language processing, and robotics. From the beginning, what might be called the purely academic interest of comprehending intelligence – and (re-)creating it within a machine – has driven many research and development goals. Early research centers pursuing this goal include MIT, Carnegie Mellon, and Stanford, all of which began their work on AI in the 1960s. However, the interests of key funders included important practical applications to balance a more theoretical perspective. For example, substantial early funding for AI research was granted by the Defense Advanced Research Projects Agency (DARPA)'s Information Processing Techniques Office, which considered AI vital to the development of technologies like advanced command-and-control systems.[2]

In 1966, John McCarthy hosted a series of four simultaneous computer chess matches carried out via telegraph against rivals in Russia.

In 1978, the applied mathematician Richard Bellman defined AI as the automation of activities associated with human thinking, which include decision-making, problem solving, and learning.[3] Critically, Bellman's definition did not emphazie a comprehensive intelligent system, but instead outlined what different components of intelligence entail. Despite early gains in areas like heuristic search, funding organizations became disillusioned by the slow pace of advancement in the 1980s, a period that became colloquially known as the "AI winter." Interest revived with the continued sophistication of computer hardware and software, as faster computers brought complex AI calculations into a range of practical utility, and in roughly the last fifteen years commercial applications for those AI techniques were found. Computer vision has enabled automated diagnostics for medical imaging. Machine learning supports predictive business analytics as well as robotics with industrial, service-based, and domestic applications. In general, many current applications entail the automation of data collection, classification, and analysis with goals of increasing efficiency and cutting costs.

The challenge of achieving artificial general intelligence (AGI), also referred to as "strong AI," depends upon the unresolved challenges of understanding and measuring human intelligence. One early benchmark for assessing successful AGI was the Turing Test, which suggests that if a human conversing with a machine cannot distinguish that machine from a human conversationalist, the machine had effectively "passed" as a human.[4] This benchmark has been criticized because AI can now pass the Turing Test by manipulating the symbols that underpin language, while still failing to understand the full meaning of dialogue. This exemplifies how humans and machines "think" in fundamentally different ways. Rather than replicating human ways of thinking, superintelligence may ultimately be achieved through human computation,[5] an approach that bring together complementary human and machine intelligence in integrated systems.

A recent analysis of news archives found a sharp increase in coverage of AI beginning in 2009.[6] As in classic science fiction writing of the 1950s and 1960s, current media hyperbole still exaggerates near term capabilities. Recent interest in AI does not seem to coincide with new fundamental advancements in the field. At the same time, the steady increase in computing speed (following Moore's Law)[7] has finally made it possible to implement practical applications of existing AI technologies.

As in classic science fiction writing of the 1950s and 1960s, current media hyperbole still exaggerates near term capabilities.

Scientific advances here and abroad, commercialization, and social response have led to increased attention to AI in numerous policy spheres. In the United States (U.S.), the first policy guidance related specifically to AI was issued by the Obama Administration. Together, a report on *Preparing for the Future of Artificial Intelligence*[8] and companion *National Artificial Intelligence Research and Development Strategic Plan*[9] offer an overview of recent AI R&D, discuss implications for fairness, accountability, safety, and security, and advance recommendations for research, regulation, and future policymaking. A third report from the Council of Economic Advisors specifically addressed the importance of AI to the American workforce, noting that "*aggressive policy action will be needed to help Americans…to ensure that the enormous benefits of AI and automation are developed by and available to all.*"[10] A flurry of high-level policy guidance was also issued by China, beginning with the *Made in China 2025* policy roadmap smart manufacturing and culminating in a high-level roadmap for becoming a global innovation center by 2030.[11] As AI increasingly becomes a topic in global policy discussions, it becomes increasingly important to understand how the benefits and risks of AI can unfold in practice, and set near- and long-term priorities accordingly.

Case study of the Research, Development, and Policy context: Autonomous Vehicles

The 2016 Transportation Statistics Annual Report, published by the Bureau of Transportation Statistics, counts 2015 transportation related fatalities at 37,000. The vast majority of these (95%) were highway fatalities, with 70% of single vehicle fatal crashes and 50% of multi-vehicle fatal crashes partially or entirely caused by human factors.[12]

By circumventing human vulnerabilities to distraction or intoxication, autonomous vehicles can reduce these fatalities. Autonomous vehicles also have a number of auxiliary advantages.[13] They can increase the quality of life for populations with limited mobility while providing drivers more time to engage in other activities. Through coordination and real-time data processing, autonomous vehicles can find the most efficient pathways for individual cars leading to decreases in congestion. More efficient driving patterns also lead to increased fuel efficiency. When combined with the increasing use of alternative energy sources, this leads to cheaper fuel prices that benefit consumers and lower fossil fuel emissions that benefit the environment.

Illustration of future roadways.

Achieving these benefits is the motivation behind new legislation. H.R. 338, the SELF- DRIVE Act, was passed in the U.S. House of Representatives on September 7, 2017. In the Senate, the AV-START Act was previously introduced as a bipartisan bill by Senators John Thune – Chairman of the Committee on Commerce, Science, and Transportation – and Gary Peters. Recognizing key differences between autonomous and human drivers, one aim of this bill is to reduce restrictive federal regulations to avoid stalling research and development.

Beyond legislation there are numerous incentives driving autonomous vehicles R&D. Federal agencies such as the National Science Foundation (NSF), the Department of Defense (DoD), and DARPA fund basic and applied research. Agencies also establish competitions, including the DARPA Grand Challenge, the Second Grand Challenge, and the Urban Challenge since 2004.[14] In addition, the Department of Transportation (DoT) initiated a $40 million competition for medium-sized cities to imagine smart city infrastructure for transportation in 2016. And while many AI techniques were historically developed in research universities, companies like Uber increasingly lure recent graduates, faculty, and even students from universities like Carnegie Mellon and the University of Washington. These tech companies are incentivized by investments from vehicle manufacturers, such as Ford Motors' recent $1 billion investment in the startup Argo AI.[15]

There are some limitations to autonomous vehicle technology. Early testing has incurred fatalities and property damage, and the ethical considerations of fault and responsibility are open and important questions. The SELF- DRIVE Act addresses safety considerations by stipulating that the Secretary of Transportation must issue a final rule regarding safety certifications within 24 months of passage. The current auto manufacturing industry, as well as related transportation segments such as public transit and crash repair, will be impacted in various anticipated and unknown ways.

Nissan Corporation has taken a hybrid approach to the driverless car problem that involves combining autonomous vehicles with a human call center. When a vehicle encounters a situation that isn't navigable, it sends telemetry to the call center where a human worker interprets the situation and provides remote guidance to the vehicle. This example of human computation illustrates how a partnerships between autonomous and humans intelligence can enable a more robust solution to safety issues. Further, while human computation systems are within easy reach, fully machine-based solutions to autonomous driving could still take decades.

> Early testing has incurred fatalities and property damage, and the ethical considerations of fault and responsibility are open and important questions.

What are the benefits of AI?

The benefits to future artificial intelligence include the ability of AI advances to augment or solve problems related to human limitations, to improve the human condition, and to build capacity to meet complex problems.

Solving problems related to human limitations. Boredom and fatigue can interfere with human performance in ways that can be addressed by automation through robotics. While currently most concentrated in auto manufacturing, smart manufacturing has the potential to benefit every industry—especially when the targets of manipulation are microscopic or smaller, as in nanotechnology. AI also has the potential to support human safety, for example through the development of autonomous vehicles (see case study).

AI can also augment human capacity where big data is concerned, by collecting, classifying, and cataloging information faster and better than human cognition alone. In automating those processes there are opportunities to gain insights about patterns and correlations that would not otherwise be feasible. The potential for automation to support social good is particularly notable in healthcare, where AI can help researchers understand biological complexity and design appropriate interventions, for example through precision medicine.[16] Looking forward, AI will continue to support healthcare advances as predictive models from machine learning support processes like triaging, and robots are able to analyze samples with greater speed and delicacy than human capabilities can offer.[17]

Improving the human condition. As AI advances permeate a range of work and leisure activities and increasingly converge with other new and emerging technologies, benefits to the human condition can begin to scale. Simple workplace automation leads to efficiency gains and a net decrease of the quantity of overall work required. These gains simultaneously increase the quality of work by liberating human energy for creative rather than menial activities. And outside of the workplace AI technologies can help meet human needs for interpersonal connection and self-actualization. Social robots developed to support Eldercare and creativity support tools bring added social and cultural value to an increasingly automated society. The potential for AI to permeate a range of life contexts through cascading effects is a benchmark of technologies ushering in the "fourth industrial revolution" (4IR), where disruptions are felt across a range of life contexts and domains as physical, digital, and biological systems merge.[18]

Building capacity to meet complex, wicked problems. Many technological inventions that reach the scope and scale of AI can create secondary problems. Medical inventions increase lifespans, and over time may exacerbate overpopulation. Consequent increases in population give viruses more opportunities to recombine into deadly strains. The invention of modern air travel allows these strains to be quickly transported all over the world. Thus, averting catastrophes at scale requires the ability to model and reason about complex interactions and dependencies, such as secondary effects.[19] By modeling and predicting complex interdependent systems, AI can help provide decision support for various futures scenarios. For example, work is underway to develop domain ontologies and automated processes for helping emergency managers and first response teams make critical, high-pressure decisions with minimal situational information.[20]

What are the risks of AI?

Despite the promise of artificial intelligence, researchers and the public at large worry about short- and long-term economic risks, such as job loss, and risks to human safety and security. Researchers also increasingly point to the social risks of AI, such as the potential for biased algorithms to cause wide-scale discrimination and perpetuate existing inequalities.[21]

Economic risks. As with other technological revolutions, artificial intelligence will disrupt the labor force by changing the work tasks performed and altering employment opportunities. According to a PEW Research public opinion study, 65% of Americans believe that robots or computers will do much of the work currently conducted by humans over the next 50 years.[22] Expert predictions vary widely. By examining the potential automation of work *tasks,* one panel estimated that 47% of jobs are automatable within the next few decades.[23] By emphasizing the heterogeneity of tasks in any given occupation, a second study found that only 6% of jobs are truly vulnerable, as the nature of many positions may change over time.[24] Due to the disparity between public and different expert opinions, some policy guidance emphasizes preparing the human workforce to better compliment machine intelligence, like through worker education.[25] Indeed, human computation systems (such as Nissan's driverless car call center) are creating new kinds of employment opportunities today. Further, just as the industrial revolution replaced the drudgery of repetitive factory jobs with new engineering jobs, today's cognitive revolution will lead to more jobs that require uniquely human abstract reasoning, creativity, and the application of world knowledge.

Risks to human safety and agency. Prominent thinkers, including Elon Musk and Nick Bostrom, caution that autonomous systems will learn to alter the course of their evolution, overriding the constraints of their human creators. That moment when AI sentience becomes superhuman is termed the "singularity." Although singularity could pose an existential threat to humanity, by and large the scientific community believes that the benefits of continuing AI R&D outweigh the risks. One recent survey found that while 31% of scientists believe that super-intelligence would be "on balance" or "extremely" bad for humanity, 52% believe that these developments will be "on balance" or "extremely" good.[26]

A more immediate risk involves the growing use of remote-controlled weapons ("drones") in war fighting. Although drones are not fully autonomous – they do not choose targets – they can navigate by GPS following a preprogrammed route. Drones avoid unnecessarily putting the lives of U.S. soldiers at risk, and can navigate environments that are difficult for human warfighters to traverse. However, the application of existing international law such as the Geneva Conventions to autonomous weapons remains unclear. One further concern is that the absence of international agreements restricting the development and use of autonomous weapons might lead to a global arms race.

In another case, today's surveillance systems help law enforcement solve crimes through video evidence. Advances in pattern recognition, enabled by AI research, allow law enforcers to detect criminal activities as they occur. The capacity for automated surveillance may challenge security through individual and cultural notions of privacy and autonomy. Due to potential biases in training, there are additional risks to surveillance, as alluded to below.

> In cases where ample and comprehensive data exist, many algorithms are sensitive enough to triangulate confidential information, for example by linking disparate pieces of information such as name and birthdate, employment history, and the presence of a chronic disease. This could facilitate discrimination by institutions ranging from banks to insurance and healthcare providers, who could deny services to "high-risk" individuals on dubious ethical or legal grounds.

Social risks. Many of the social risks associated with AI derive from the combination of big data and automated decision making. Fueling concerns in this area is a recent boom in machine learning (ML), a technique focused on "teaching" algorithms to make predictions about the future by analyzing data about relevant present and past conditions. In cases where ample and comprehensive data exist, many algorithms are sensitive enough to triangulate confidential information, for example by linking disparate pieces of information such as name and birthdate, employment history, and the presence of a chronic disease. This could facilitate discrimination by institutions ranging from banks to insurance and healthcare providers, who could deny services to "high-risk" individuals on dubious ethical or legal grounds. Because algorithmic methods for decision-making in these contexts are often protected as trade secrets, such discrimination would be difficult to prove.[27]

But data sets are more often than not incomplete, either lacking or over-representing information from historically marginalized groups. This has resulted in software with a range of biases. One study exploring the "blackbox" of algorithms found that advertisements for high-paying jobs were more likely to target men than women.[28] Other research finds that the identification abilities of facial recognition software depend on diverse training data.[29] Algorithms are trained on datasets that are evenly distributed across all demographics, software will equally recognize all racial cohorts. Training data that represents existing biases—such as the overrepresentation of African Americans in criminal justice systems, and therefore criminal mug shots—will lead to automated systems that reinforce these biases by creating algorithms that are more likely to, for example, associate African American faces with crime.

Planning in an Age of Complexity: Recommendations for Policymakers and Funders

AI is a critical component of the fourth industrial revolution (4IR), "a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres."[30] Compared to previous revolutions involving *processes* like mechanization, mass production, and automation, the fourth industrial revolution is characterized by the *convergence* of new and emerging technologies in, complex socio-technical systems that permeate every aspect of human life.[31]

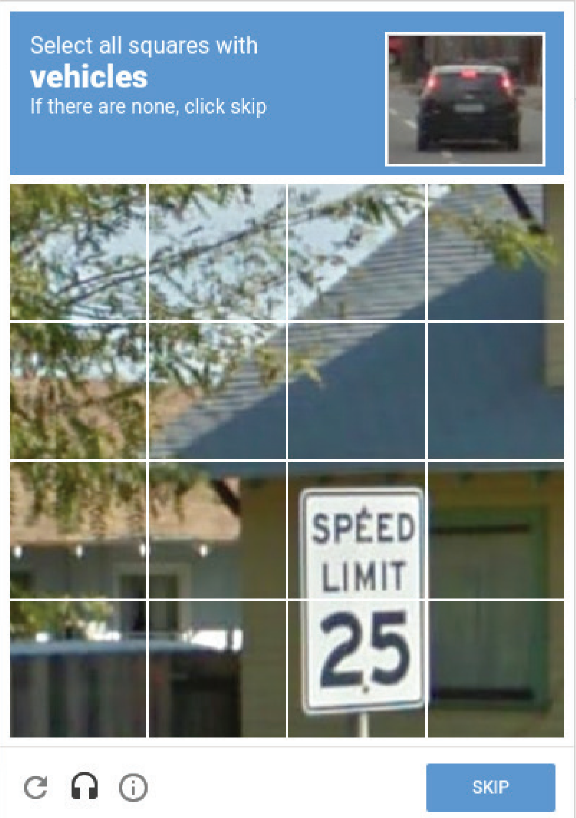
Convergence also implies the increasing interaction of multiple fields, such as AI, genomics and nanotechnology, which rapidly expands the range of possible impacts that need to be considered in any science policy exercise.[32] Ten years ago, nanotechnology was celebrated largely for its impacts on chemistry and material sciences. But the ability to precision engineer matter at genetically relevant scale has resulted in significant advances in genomics and neurosciences, such as creating the ability to model networks of neurons.[33] This example illustrates how the convergence of two emerging technologies —AI and genomics—leads to advances beyond the initial capabilities of either alone. Meeting the challenges of convergence requires drawing on a wide range of expertise, and taking a systems approach to promoting responsible research and innovation.[34]

As "outsiders" to the AI design processes, it is extremely difficult for policymakers to estimate AI development due to limited comprehension of how the technology functions. Many may also draw inspiration from traditional regulatory models that are inadequate for AI, playing a catch-up game to decode the terms of reference used by researchers, or fall victim to the human fallacy of overestimating the short-term capabilities of new technologies. There will be significant systems' transformations through AI over the next few decades, but perhaps it will be more incremental than we fear or imagine.

Recommendation for policymaking and advisory bodies: Conduct broad and deep investigations into AI with leading researchers from the private sector and universities.

In the US, early reports from policy bodies and researchers at institutions such as Stanford[35] offer high-level roadmaps of AI R&D. Expert groups convening under organizations like IEEE compliment these overviews with in-depth considerations of things like ethically-aligned AI design to maximize human well-being.[36] In the near-future, AI researchers involved in collaboration with policymakers should conduct additional in-depth studies to better understand and anticipate aspects of AI related to (for example) job automation at a more granular level, considering impact across time, sectors, wage levels, education degrees, job types and regions. For instance, rather than low-skills jobs that require advanced hand-dexterity, AI systems might more likely replace routine but high-level cognitive skills. Additional studies could investigate areas like national security.

Recommendation for policymaking and advisory bodies: Advocate for a systems approach to AI research and development that accounts for other emerging technologies and promotes human participation.



reCAPTCHA is an example of human computation.

AI seeks to replicate human intelligence in machines – but humanlike intelligence already exists in humans. Today there is an opportunity to develop superhuman intelligence by pairing the complementary abilities of human cognition with the best available AI methods to create hybrid distributed intelligent systems. In other words, it is in our reach to build networks of humans and machines that sense, think, and act collectively[37] with greater efficacy than either humans or AI systems alone. The emerging subfield of AI known as Human Computation[38] is exploring exactly those opportunities by inserting humans into the loop in various information processing systems to perform the tasks that exceed the abilities of machine AI. For this reason, human computation is jokingly referred to as "Artificial AI".

Powerful examples of human computation can be found in citizen science platforms, such as stardust@home, in which 30,000 participants discovered five interstellar dust particles hidden among a million digital images, or Stall Catchers, where thousands of online gamers help accelerate a cure for Alzheimer's disease. Human Computation also includes crowdsourcing platforms like Wikipedia, which aggregates information into the world's most comprehensive collection of human knowledge, or reCAPTCHA, which has the dual purpose of validating humans for websites while extracting information from images. Furthermore, platforms like Habitat Network leverage the reciprocity of online and offline behavior to influence conservation behavior in people's own backyards. This kind of collective action on a massive scale may be the only foil against societal issues that arose due to deleterious collective action, such as combustion engine usage. Examples from these "first generation" human computation applications only hint at the untapped potential of hybrid human/machine systems.

Human computation approaches to AI address such dystopian concerns by advancing the design of AI systems with human stakeholders in the loop who drive the societally-relevant decisions and behaviors of the system. The conscientious development of AI systems that carefully considers the coevolution of humans and technology in hybrid thinking systems will help ensure that humans remain ultimately in control, individually or collectively, as systems achieve superhuman capabilities.

Recommendation for policymaking and advisory bodies: To promote innovation, avoid centralizing and dramatically expanding regulation.

Current regulations regarding AI are as additions to products, and thus, subsumed under existing legislative authorities. For example, the Food and Drug Administration (FDA) regulates precision medicine initiatives, while the National Highway Traffic Safety Administration under the DoT issues guidance around autonomous vehicles while leaving key decisions to the states. Experts agree that expanding and centralizing regulation will inhibit growth. Recognizing that "ambiguous goals" coupled with broad transparency requirements have encouraged firms to prioritize consumer privacy, some researchers advocate for a similar approach in AI, where strong transparency requirements promote innovation yet force accountability.[39]

At the same time there is a need to safeguard the public interest by addressing shared challenges. That includes developing safety guidelines and accident investigation processes, protecting individual privacy and public security, designing systems with transparent decision-making, and managing public perception through effective science communication.

Recommendation for policymaking and advisory bodies: Seek to maximize benefits and minimize risks by issuing guidance and by building or supporting subject-specific communities. These efforts, in turn, help to encourage broad perspective and creative activity.

To promote top-down coordination within the federal government, the Obama Administration chartered a National Science and Technology Council (NSTC) Subcommittee on Machine Learning and Artificial Intelligence in 2016. But cross-agency innovation also comes from grassroots, bottom-up communities. The federal communities of practices (COP)s supported by Digital.gov are important hubs for sharing and developing common strategies. One COP already explores the role of artificial intelligence in delivering citizen services.[40] Executive guidance could encourage more COPs to form.

The public sector is an important stakeholder in AI R&D, as well as a valuable source of expertise. As a first step towards engaging the public the Obama Administration co-hosted five workshops with a few universities and NGOs and issued a public Request for Information (RFI).[41] Additional gatherings could leverage the knowledge of industry leaders as well as state and local government advocates and customers. Beyond one-off events, in line with government ideals of openness and transparency ongoing programs for public engagement and outreach should also be supported. One example of a program already active is the Expert & Citizen Assessment of Science & Technology (ECAST) network,[42] which brings universities, nonpartisan policy research organizations, and science educators together to inform citizens about emerging science and technology policy issues, and solicit their input to promote more fully informed decision-making.

Recommendation for funders: Advance artificial intelligence R&D through grants as well as prize and challenge competitions.

While basic research can take decades to yield commercial results, research such as Claude Shannon's information theory,[43] which enabled later developments such as the Internet, illustrate the value of theoretical work. Foundational research to advance AI should help develop safety and security guidelines, explore ways to train a capable workforce, promote the use of standards and benchmarks, and unpack key ethical, legal, and social implications. Agencies should continue to prioritize areas including robotics, data analysis, and perception while also advancing knowledge and practice around human-AI collaboration, for example through human computation.

Applied AI R&D unfolds in an innovation ecosystem where diverse stakeholders strategically collaborate and compete by applying for funding, procuring new technologies, and hiring top researchers. Funding agencies should incentivize collaboration by bringing industry, academia, and government interests together in pre-competitive research. NSF's Industry–University Cooperative Research Centers Program (IUCRC) program already enables such collaboration to advance foundations for autonomous airborne robotics. There are also Small Business Innovation Awards (SBIRs), which can allow small companies the opportunity to explore high-risk, high-reward technologies with commercial viability.

While the impacts of prize and challenge competitions are difficult to quantify, DARPA believes that ten years after the Grand Challenge "*defense and commercial applications are proliferating*."[44] Prizes and challenges allow agencies to incentivize private sector individuals and groups to share novel ideas or create technical and scientific solutions. These mechanisms are low-cost opportunities to explore targeted AI developments that advance specific agency missions.

ENDNOTES

1 Rajaraman, V. Reson. 2014. John McCarthy—Father of Artificial Intelligence. Resonance 19 (No.3): 198-207.

2 National Research Council. 1999. Funding a Revolution: Government Support for Computing Research. Washington, DC: The National Academies Press. https://doi.org/10.17226/6323.

3 Bellman, Richard. 1978. Artificial Intelligence: Can Computers Think?. Thomson Course Technology.

4 Turing, Alan Mathison. 1937. On Computable Numbers, with an Application to the Entscheidungsproblem, 230-265. Proceedings of the London Mathematical Society.

5 Michelucci, Pietro. 2013. "Human Computation: A Manifesto." Handbook of Human Computation. Springer, New York.

6 Fast, Ethan and Eric Horvitz. 2016. "Long-Term Trends in the Public Perception of Artificial Intelligence." http://arxiv.org/abs/1609.04904.

7 Waldrop, M. Mitchell. 2016. "The chips are down for Moore's law." http://www.nature.com/news/the-chips-are-down-for-moore-s-law-1.19338.

8 United States. Executive Office of the President. National Science and Technology Council. Committee on Technology. Preparing for the Future of Artificial Intelligence. Washington, DC: United States National Science and Technology Council, Committee on Technology, 2016. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

9 United States. Executive Office of the President. National Science and Technology Council. Networking and Information Technology Research and Development Subcommittee. The National Artificial Intelligence Research and Development Strategic Plan. Washington, DC: United States National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, October 2016. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf.

10 United States. Executive Office of the President. Artificial Intelligence, Automation, and the

Economy. Washington, DC: United States Executive Office of the President, December 2016. https://www.whitehouse.gov/sites/whitehouse.gov/files/images/EMBARGOED%20AI%20Economy%20Report.pdf.

11 China. State Council. China issues guideline on artificial intelligence development, 2017. http://english.gov.cn/policies/latest_releases/2017/07/20/content_281475742458322.htm.

12 United States. Department of Transportation. Bureau of Transportation Statistics. Transportation Statistics Annual Report 2016. Washington, DC: 2016. Chapter 6. https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/transportation_statistics_annual_report/index.html.

13 Anderson, James M., Nidhi Kalra, Karlyn Stanley, Paul Sorensen, Constantine Samaras and Tobi A. Oluwatola. Autonomous Vehicle Technology: A Guide for Policymakers. Santa Monica, CA: RAND Corporation, 2016. http://www.rand.org/pubs/research_reports/RR443-2.html.

14 United States. Department of Defense. Defense Advanced Research Projects Agency. The DARPA Grand Challenge: Ten Years Later. Washington, DC: March 2014. https://www.darpa.mil/news-events/2014-03-13.

15 Forbes. 2017. "Here's Why Ford Is Investing In Argo AI." https://www.forbes.com/sites/greatspeculations/2017/02/27/heres-why-ford-is-investing-in-argo-ai/#1e96fd56250a.

16 Pauwels E. and A. Vidyarthi. 2017. Who Will Own The Secrets in Our Genes? A U.S.-China Race in Artificial Intelligence and Genomics, Policy brief, Wilson Center. https://www.wilsoncenter.org/sites/default/files/who_will_own_the_secrets_in_our_genes.pdf

17 Pauwels E. and Garrett Dunlap. 2017. "The Intelligent and Connected Bio-Labs of the Future: Promise and Peril in the Fourth Industrial Revolution." World Economic Forum. https://www.weforum.org/agenda/2017/10/intelligent-and-connected-bio-labs-promise-and-peril-in-the-fourth-industrial-revolution

18 Pauwels E. and Garrett Dunlap. 2017. "The Intelligent and Connected Bio-Labs of the Future: Promise and Peril in the Fourth Industrial Revolution." World Economic Forum.

19 Pauwels E., Rejeski D., and J. Koo. 2016. Science and Technology Forecasting, W.S. Bainbridge, M.C. Roco (eds.), Handbook of Science and Technology Convergence.

20 Haghighi, Pari Delir, Frada Burstein, Arkady Zaslavsky, and Paul Arbon. 2013. "Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings." Decision Support Systems 54, no. 2 (January): 1192-1204. http://www.sciencedirect.com/science/article/pii/S0167923612003454.

21 Crawford, Kate. 2016. "Artificial Intelligence's White Guy Problem." New York Times, June 25, 2016. https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html.

22 Smith, Aaron. 2016. "Public Predictions for the future of Workforce Automation." Pew Research Center. March 10, 2016. http://www.pewinternet.org/2016/03/10/public-predictions-for-the-future-of-workforce-automation/.

23 Frey, Carl Benedikt and Michael A. Osborne. "The Future of Employment: How Susceptible are Jobs to Computerisation?" Oxford University. 2013. https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf.

24 Arntz, M., T. Gregory and U. Zierahn. "The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis." OECD Social, Employment and Migration Working Papers, No. 189, OECD Publishing. 2016. http://dx.doi.org/10.1787/5jlz9h56dvq7-en.

25 Council for Economic Advisors, 2016.

26 Müller Vincent C. and Nick Bostrom. "Future Progress in Artificial Intelligence: A Survey of Expert Opinion." Fundamental Issues of Artificial Intelligence. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science), vol 376. 2016. https://doi.org/10.1007/978-3-319-26485-1_33.

27 Pauwels E. and A. Vidyarthi. 2017. Who Will Own The Secrets in Our Genes? A U.S.-China Race in Artificial Intelligence and Genomics, Policy brief, Wilson Center. https://www.wilsoncenter.org/sites/default/files/who_will_own_the_secrets_in_our_genes.pdf

28 Datta, Amit, Michael Carl Tschantz and Anupam Datta. 2015. "Automated experiments on Ad privacy settings." Proceedings on Privacy Enhancing Technologies, 1, 92-112.

29 Klare, Brendan F., Mark J. Burge, Joshua C. Klontz, Richard W. Vorder Bruegge and Anil K. Jain. 2012. "Face recognition performance: Role of demographic information." IEEE Transactions on Information Forensics and Security, 7, 6, 1789-1801.

30 Schwab, Klaus. 2016. "The Fourth Industrial Revolution: what it means, how to respond." World Economic Forum. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond.

31 Pauwels E., Rejeski D., and J. Koo. 2016. Science and Technology Forecasting, in W.S. Bainbridge, M.C. Roco (eds.), Handbook of Science and Technology Convergence.

32 Pauwels E. 2017. The Great US-China Biotechnology and Artifical Intelligence Race, The Diplomat. https://thediplomat.com/2017/08/the-great-us-china-biotechnology-and-artificial-intelligence-race/

33 Pauwels E. and W. Johnson. 2017. How to Optimize Human Biology: Where Genome Editing and Artificial Intelligence Collide, World Economic Forum. https://www.weforum.org/agenda/2017/11/how-to-optimize-human-biology

34 Pauwels E. 2011. "Who Let the Humanists into the Lab?," Valparaiso University Law Review's 2011 Symposium Issue –The Synthetic Cell: Bioethics, Law and the Implications of Synthetic Life, Volume 45.

35 Stanford University, accessed August 1, 2016. "One Hundred Year Study on Artificial Intelligence (AI100)." https://ai100.stanford.edu.

36 IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems. 2016. "Ethically Aligned Design." http://standards.ieee.org/develop/indconn/ec/ead_v1.pdf.

37 Michelucci, Pietro. 2013. "Organismic Computing." Handbook of Human Computation. Springer New York, pp. 475–501.

38 "Human Computation: A Manifesto," Michelucci.

39 Stone, Peter, et al. "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel. Stanford: Stanford University 2016. Accessed February 8, 2017. https://ai100.stanford.edu/2016-report.

40 DigitalGov. n.d. "Communities of Practice." Accessed November 9, 2017. https://www.digitalgov.gov/communities/.

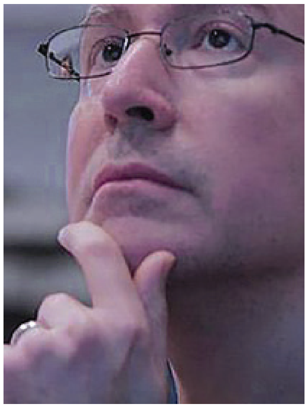
41 National Science and Technology Council, Preparing for the Future of Artificial Intelligence.

42 Expert & Citizen Assessment of Science & Technology. n.d. "A Distributed Network of Institutions for Peer to Peer Public Deliberation." https://ecastnetwork.org/.

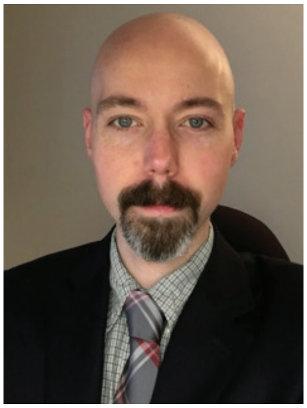
43 Shannon, C.E. 1948. "A Mathematical Theory of Communication." Bell System Technical Journal, 27, pp. 379–423 & 623–656. July & October, 1948.

44 DARPA, The DARPA Grand Challenge: Ten Years Later.

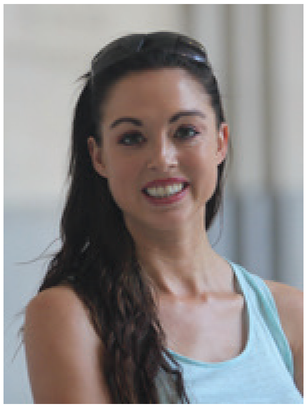
Dr. Anne Bowser is Director of Innovation with the Science and Technology Innovation Program at the Wilson Center. In the context of artificial intelligence and other emerging technologies, Bowser is interested in developing processes for involving the public in research and development, like through crowdsourcing and human computation, and in policy dialogues. Expanding governance around emerging technologies can accelerate discovery while enhancing public understanding of science.



Dr. Pietro Michelucci is Director of the Human Computation Institute, a multidisciplinary innovation center that develops crowd-powered systems to address societal problems. He is Editor-in-Chief of the 2013 Springer Handbook of Human Computation and Founding Editor of the journal Human Computation. As an Organismic Computing pioneer, he is interested in developing new methods for augmenting collective intelligence. In 2015 Michelucci launched the EyesOnALZ citizen science project to accelerate Alzheimer's disease research.



Michael Sloan is an analyst in technology policy. He specializes in research and development programs, and contributed to this publication while working as a Research Assistant with the Science and Technology Innovation Program at the Wilson Center. Sloan graduated with an M.A. in Science and Technology Policy from George Washington University in 2017.



Eleonore Pauwels is the Director of the AI Lab with the Science and Technology Innovation Program at the Wilson Center. Pauwels is a writer and international science policy expert, who specializes in the governance and democratization of converging technologies. Leading the AI Lab, Pauwels analyzes and compares how transformative technologies, such as artificial intelligence and genome-editing, raise new opportunities and challenges for health, security, economics and governance in different geo-political contexts.

The Wilson Center

@TheWilsonCenter

facebook.com/WoodrowWilsonCenter

www.wilsoncenter.org

Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue NW
Washington, DC 20004-3027

Wilson Center