Creation of the National Artificial Intelligence Research and Development Strategic Plan

Lynne E. Parker

■ *In October 2016, the United States government published the National Artificial Intelligence Research and Development Strategic Plan, which lays out a strategic plan for federally funded research and development in AI. As a coleader of the task force that developed this plan, I was asked to discuss its creation in an invited presentation at AAAI 2017. This article is based on that presentation, which outlines not only the plan itself, but also provides insight into its goals and objectives, and background into how it was created.*

The significant attention paid to AI in the popular press in recent years has led to growing uncertainties among nonexperts in AI about what AI can and cannot do, and what the consequences are of ever more capable AI. The general public wonders whether AI is going to take away their jobs or maybe even take over the world. Some believe that, given the success of deep learning for some industrial applications, AI must now be solved, or at least advanced to the point that industry can now address all remaining challenges. Given the substantial growth of industry around AI products, there has also been uncertainty on what the continued role of the government should be in AI investments.

In 2016, amidst this landscape of uncertainty, the United States government launched a series of activities and actions to help the country better understand and prepare for the impacts of advancements in artificial intelligence (Felten 2016b). In one of those actions, the government called for the creation of a national strategic plan on AI that defines the federal role in AI research and development (R&D). Why was a strategic plan needed? If done thoughtfully, a national AI R&D strategic plan could help address these uncertainties by identifying the federal role in AI investments and defining open AI R&D challenges that must be solved before AI can be used in important societal applications. Such a plan could provide structure to the field in terms of current AI capabili-

ties, desired future AI capabilities, and how best to focus R&D investments to achieve the desired future.

To develop this plan, an interagency task force of the federal government was created in the summer of 2016, leading to the release of the National AI Research and Development Strategic Plan[1] on October 12, 2016 (Felten and Lyons 2016a). The goal in putting together this plan was to create a high-level framework to identify scientific and technological needs in AI over the next 5 to 15 years. Particular attention was paid to (1) longer-term transformational impacts of AI, (2) the role of the federal government in AI investments, and (3) AI challenges that are unlikely to be addressed by industry.

Why these three areas of emphasis? First, it is recognized that many of today's significant industry successes in AI build upon pioneering transformational ideas developed many years ago by early investigators in the field. Much of this pioneering research was funded by the federal government, along with important industry investments in basic research laboratories (National Research Council 2012). Today, the funding landscape has evolved, such that the federal government is the primary source of funding for long-term, high-risk research initiatives, as well as near-term research for agency-specific requirements that industry does not pursue. This is not to say that industry is not continuing R&D investments in AI research; indeed, important AI advances are being made by industry-based researchers. However, much of today's industrial research is aimed at consumer markets and near-term profit drivers. The National AI R&D Strategic Plan instead focuses on those longer-term and mission-focused investments that are unlikely to be duplicated by industry.

Who Are the Intended Audiences for the Plan?

The primary intended audiences for the National AI R&D Strategic Plan are the US policymakers and federal funding agencies who support research and development in AI. While the plan does not stipulate specific funding programs for individual agencies, it does give a broad perspective on high-priority funding areas in AI for the federal government as a whole. In addition to informing policymakers and federal funding agencies, this strategic plan has also served an important role in organizing some of the public dialog around AI. For example, together with other important reports such as Preparing for the Future of Artificial Intelligence[2] and the 2016 report from the One Hundred Year Study on Artificial Intelligence (Stone et al. 2016), it places into context the potential of AI to enhance the quality of life across a broad swath of societal areas, as well as the open challenges being faced in creating these AI solutions. Numerous public seminars, symposia, forums, summits, and workshops are considering tough questions about AI.

When these venues delve into areas of research and the government's role in AI, the plan provides useful organization and guiding principles. It also serves an important role in informing academic, industry, and government partnerships in AI, particularly in providing a clear understanding of the government's priority areas for AI R&D.

Who Created the Plan?

As a report of the United States federal government, the National AI R&D Strategic Plan was developed by the Artificial Intelligence Task Force, a United States interagency working group tasked by the Subcommittee on Networking and Information Technology Research and Development (NITRD). The NITRD Subcommittee is a body under the National Science and Technology Council (NSTC). NSTC, in turn, is part of the White House Office of Science and Technology Policy (OSTP). This tasking was done at the request of another NSTC subcommittee, the Subcommittee on Machine Learning and Artificial Intelligence, which was created in the spring of 2016. While not well known outside of the information technology R&D circles in the federal government, the NITRD Subcommittee plays a key role in coordinating federal investments in advanced networking and information technology across the federal government. NITRD was initially created by the High Performance Computing Act of 1991, and is composed of representatives from member agencies across the federal government who support information technology (IT) R&D. Over 20 federal departments, agencies, and offices are "members" of NITRD, including the National Science Foundation (NSF), the National Institute of Standards and Technology (NIST), the National Institutes of Health (NIH), the National Aeronautics and Space Administration (NASA), the Defense Advanced Research Projects Agency (DARPA), the Environmental Protection Agency (EPA), the National Institute of Justice (NIJ), the Department of Energy (DOE), the Department of Homeland Security (DHS), and the National Security Agency (NSA).

NITRD has a number of interagency working groups that have developed several IT-related strategic plans over the years, such as the National Privacy Research Strategy (July 2016), the Federal Big Data Research and Development Strategic Plan (May 2016), and the Federal Cybersecurity Research and Development Strategic Plan (February 2016). Because of the key role of the NITRD agencies in overseeing federal investments in AI, representatives from these agencies with technical expertise in AI worked collaboratively to bring the National Artificial Intelligence Research and Development Strategic Plan into fruition. The Artificial Intelligence Task Force thus consisted of myself as a coleader (representing NSF), along with Jason Matheny of IARPA (Intelligence

Advanced Research Projects Activity). The other task force members were (in alphabetical order) Milton Corn (NIH), William Ford (NIJ), Michael Garris (NIST), Steven Knox (NSA), John Launchbury (DARPA), Richard Linderman (DOD), Nikunj Oza (NASA), Robinson Pino (DOE), Gregory Shannon (OSTP), and Scott Tousley (DHS). Faisal D'Souza, with the National Coordination Office for NITRD, was the task force technical coordinator. Other federal agencies with IT interests also contributed to the plan. In the end, about fifteen federal agencies provided input and consensus approval for the plan's content.

What US National Priorities Guided the Creation of the Plan?

A prerequisite to the development of any strategic plan is an understanding of the high-level priorities that guide the creation of that plan. In the context of the National AI R&D Strategic Plan, what are these priorities? There are many perspectives on what a nation's priorities should be, but for the United States one can look to the founding statement on which the country was built — the Declaration of Independence — and conclude that the government is created to protect the "unalienable rights" of all human beings: "Life, Liberty, and the pursuit of Happiness," Thus, the plan's vision for advancing national priorities of AI revolve around applications that improve quality of life (for example, education, medicine), applications that enhance liberty (for example, security and safety), and applications that increase happiness, interpreted in the Plan as economic prosperity (for example, manufacturing, transportation, agriculture). Interestingly, many of these application areas align with the mission foci of various federal funding agencies.

To guide the overall vision for the plan, the task force created a vision of the future that might be possible with AI. The ultimate vision is a future world in which AI is safely used for significant benefit to all members of society, with minimal economic and societal disruption. Further progress in AI could enhance our well-being in nearly all sectors of society. We imagined a future world when investments and progress in AI can be harnessed, leading to increased economic prosperity, creating new efficiencies and new markets that drive further innovation, raising the quality of life, and strengthening national security. This vision of the future might look something like this:

> Smart farming techniques lead to high-yield, sustainable agriculture, and crops that are increasingly immune to damaging pests. Rural communities everywhere have access to top-notch healthcare through remote medicine. Epidemics are anticipated and tracked through real-time analytics, and then slowed or halted through individualized alerts and action plans. Science and innovation are advanced through intelligent data analysis, enabling better theoretical insights, concept formation, anomaly detection, and prediction of new properties in measured phenomena. Low-cost, noninvasive, mind-controlled prosthetics are available worldwide to dramatically increase mobility and independence for the disabled. Personalized education is available for everyone, enabling all to be skilled and valuable contributors to the workforce. More productive industrial processes lead to manufacturing jobs moving back onshore. Robots work alongside us to help us and care for us, giving us more years of independent living; they reduce workplace injury by taking on dangerous, dull, or dirty work. Virtual concierges use smart control systems to optimize energy use. Wearable devices and algorithms collect data and give cues to produce efficient movements of people through cities and communities, all while protecting personal privacy. Automated vehicles take us where we need to go reliably and safely, while reducing traffic and space designated for parking. Augmented and virtual reality becomes mainstream, replacing unnecessary high-carbon travel. Foolproof identity-verification technologies reduce identity theft and the need for passwords, photo identification cards, and PIN numbers. Intelligent, self-learning algorithms prevent, protect against, and defeat cyber attacks. This is a possible AI future.

While this vision was not ultimately included in the plan in the form of this narrative, it is representative of the positive outlook that the task force had on the power of AI to address our national priorities and improve our quality of life.

How Was the Content of the Plan Determined?

A key challenge in developing the National AI R&D Strategic Plan was not only to identify the key open R&D challenges of AI, but also to combine the AI perspectives of so many federal agencies into a cohesive overall strategy. Importantly, the plan is intended to be inclusive of the broad mission objectives of the variety of federal agencies who support AI R&D. For example, the mission of NSF is "to promote the progress of science; to advance the national health, prosperity, and welfare; and to secure the national defense; and for other purposes." This mission leads NSF in its vision to broadly support basic research and education, including areas of AI. The mission of NIST is "to promote US innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life." This mission translates into NIST's investments in related areas of R&D, including standards and benchmarks for AI. Similarly, each federal agency has a unique mission that covers a particular segment of national interest, which then maps to relevant focused investments in AI.

Taking these agency missions into consideration, a further goal in developing the National AI R&D
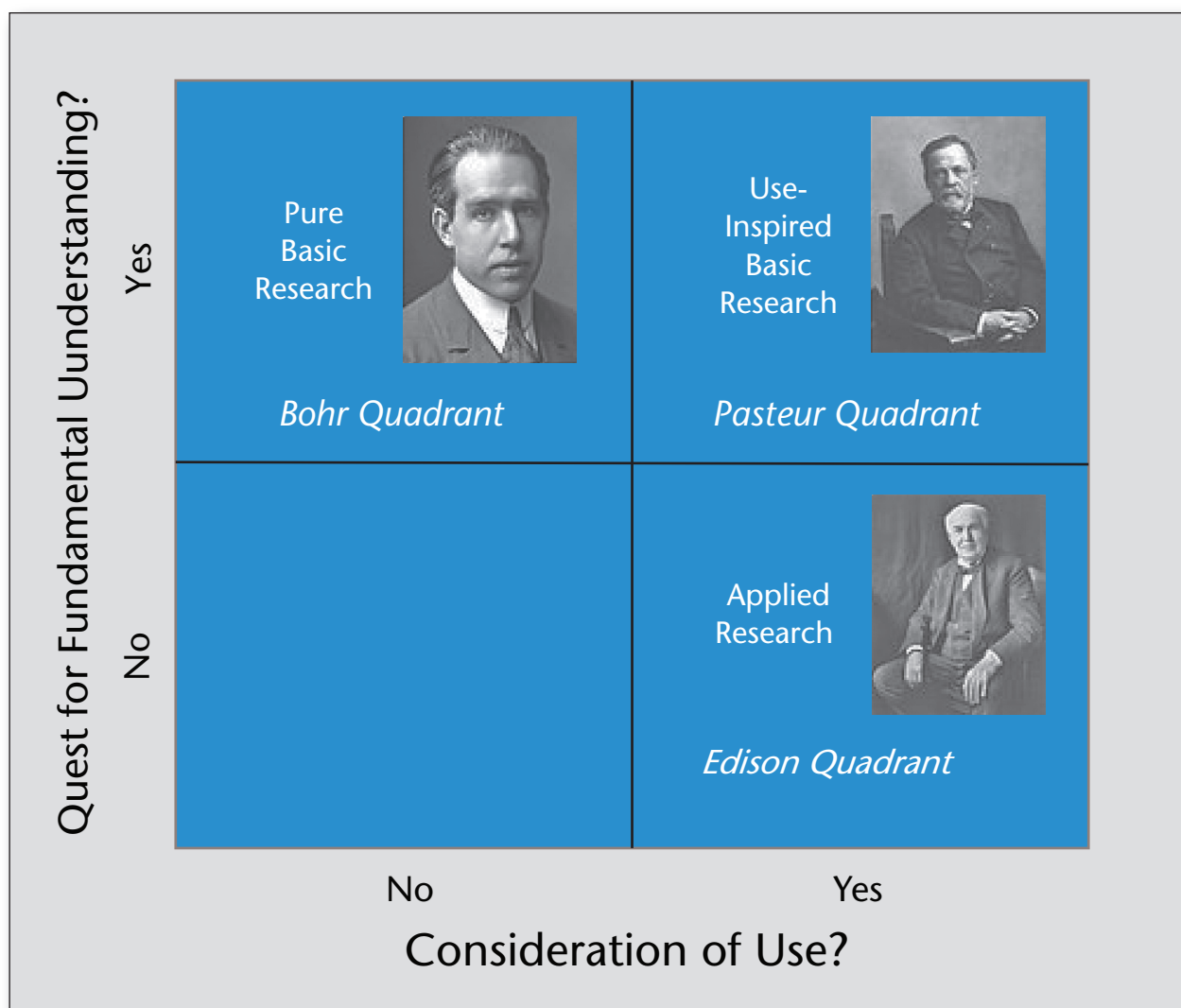
Figure 1. Classification of Research.

Research is organized into three broad categories, based on whether the research is aimed at fundamental knowledge and whether the research is done with a consideration of use in mind. The quadrant names were introduced by Donald Stokes.

Strategic Plan was to identify the most important areas of AI research that can have a broad positive impact across the entirety of national interests. Identifying this collective of R&D challenges was part of the challenge in creating the plan. Arguably, it could have been easier to catalogue each technical area of AI research and state the open R&D challenges in that area — much like organizing sessions for a conference. But an important objective in developing the plan was to put some structure on the overall research so as to emphasize the common themes that are important across multiple agencies. The plan must be representative of the priority needs and missions of the collective of federal R&D agencies as a whole, while also identifying those areas that are unlikely to receive significant industry investment.

An effective approach to defining areas of com-

mon R&D interest across multiple agencies involves outlining the priorities of each agency, finding recurring R&D themes, and defining a meaningful organization of the common R&D themes. Of course, the landscape of federal investments in AI is rather complex, since some federal agencies are focused on long-term basic research, while others focus on shorter-term mission needs. In this context, a helpful perspective was proposed by Donald Stokes' in his book *Pasteur's Quadrant* (Stokes 1997), illustrated in figure 1. In his approach, two guiding questions create three relevant quadrants of research: (1) whether the research is undertaken with a quest for fundamental understanding, and (2) whether the research is pursued with a particular use-case in mind. Longer-term, fundamental research investments without specific use cases in mind fall into the quadrant for

pure basic research, Bohr's Quadrant. Agencies such as NSF and DARPA (among others) often make investments in these areas. Pasteur's Quadrant, on the other hand, represents fundamental research applied to specific domains. NIH's interests, for example, fall into "use-inspired" basic research in the context of health and well-being. In the context of the National AI R&D Strategic Plan, Edison's Quadrant represents the AI research needs of mission agencies that have near-term, agency-specific goals that are not being addressed by industry. Parts of DoD and NIJ (among others) fund research of this type.

Aside from the federal agencies, additional input to the National AI R&D Strategic Plan came from a thorough review of the open literature on the state of the field of AI, public discussions at AI-related meetings, an Office of Management and Budget (OMB) data call across all federal agencies who invest in IT-related R&D, and a request for information (RFI) by the Office of Science and Technology Policy that solicited public opinions about how America can best prepare for an AI future (Felten and Lyons 2016b).

One additional point about the content of the plan is important to understand. Because R&D in AI primarily occurs within the discipline of information technology, the charge for the creation of the plan was directed to NITRD. Due to the fact that NITRD oversees (specifically) IT-related R&D coordination across the federal government, the content of the plan is exclusively focused on open IT-relevant issues for AI. Of course, the Artificial Intelligence Task Force recognized that AI benefits from a variety of perspectives across many other disciplines, including neuroscience, psychology, social and behavioral sciences, ethics, law, economics, as well as expertise from across the broad spectrum of application domains, including agriculture, transportation, and so forth. Research and development in these other domains is not included in the strategic plan, however, due to the IT-centric tasking of the task force. Nevertheless, a focus on IT-relevant issues still provides a useful foundation for considering priorities in AI R&D investments, and their potential benefits across a wide range of application domains.

Overview of the AI R&D Strategic Plan

Ultimately the task force defined seven strategic R&D priorities for AI that are included in the plan. While the reader is referred to the plan itself for more details of the primary areas of emphasis, a quick summary is given here for completeness, taken from the executive summary:

Strategy 1: Make long-term investments in AI research. Prioritize investments in the next generation of AI that will drive discovery and insight and enable the United States to remain a world leader in AI.

Strategy 2: Develop effective methods for human-

AI collaboration. Rather than replace humans, most AI systems will collaborate with humans to achieve optimal performance. Research is needed to create effective interactions between humans and AI systems.

Strategy 3: Understand and address the ethical, legal, and societal implications of AI. We expect AI technologies to behave according to the formal and informal norms to which we hold our fellow humans. Research is needed to understand the ethical, legal, and social implications of AI, and to develop methods for designing AI systems that align with ethical, legal, and societal goals.

Strategy 4: Ensure the safety and security of AI systems. Before AI systems are in widespread use, assurance is needed that the systems will operate safely and securely, in a controlled, well-defined, and well-understood manner. Further progress in research is needed to address this challenge of creating AI systems that are reliable, dependable, and trustworthy.

Strategy 5: Develop shared public datasets and environments for AI training and testing. The depth, quality, and accuracy of training datasets and resources significantly affect AI performance. Researchers need to develop high-quality datasets and environments and enable responsible access to high-quality datasets as well as to testing and training resources.

Strategy 6: Measure and evaluate AI technologies through standards and benchmarks. Essential to advancements in AI are standards, benchmarks, testbeds, and community engagement that guide and evaluate progress in AI. Additional research is needed to develop a broad spectrum of evaluative techniques

Strategy 7: Better understand the national AI R&D workforce needs. Advances in AI will require a strong community of AI researchers. An improved understanding of current and future R&D workforce demands in AI is needed to help ensure that sufficient AI experts are available to address the strategic R&D areas outlined in this Plan.

A concise organization of these strategies, taken from the plan, is shown in figure 2. In the bottom row of this graphic (in dark red) are the cross-cutting R&D foundations that underpin nearly all areas of AI, regardless of application. These foundational issues include ethical, legal, and societal implications, focusing on fairness, transparency, and accountability by design, as well as ethical AI (strategy 3); safety and security issues, focusing on explainability and transparency, building trust, verification and validation, and securing against attacks (strategy 4); a need for shared data sets and environments for training and testing, to accelerate the effective development of AI (strategy 5); and developing standards and benchmarks to evaluate AI systems (strategy 6).

While not a strict technical challenge, strategy 7 defines the need for a capable AI workforce for developing and using cutting-edge AI approaches. This strategy also impacts all of AI R&D.

The middle row of figure 2 (in lighter shades of blue) focuses on the basic areas of R&D that build upon the cross-cutting foundations of R&D. These
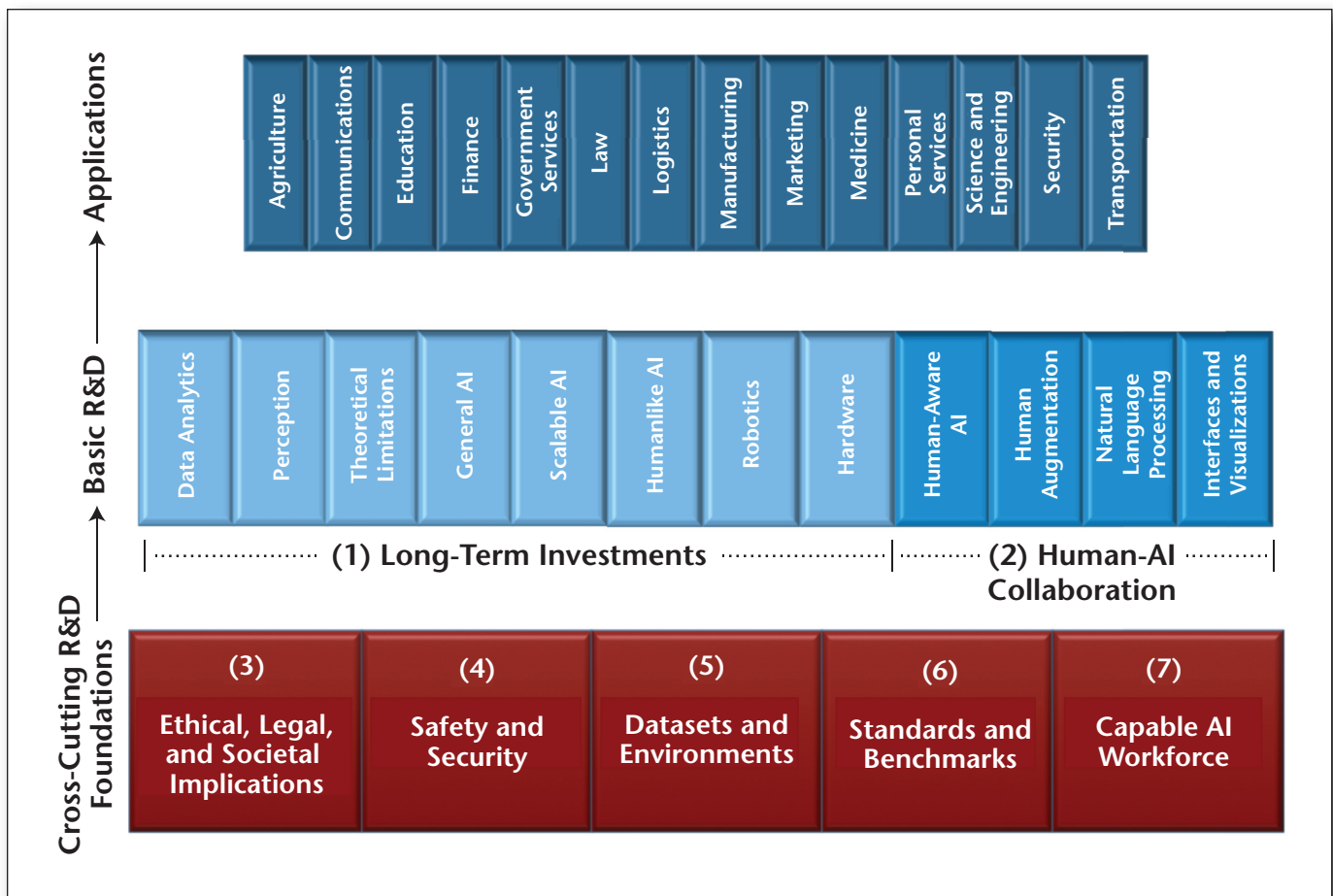
Figure 2. Organization of the R&D in the National AI R&D Strategic Plan.

Published by OSTP's NSTC.

basic research areas encompass a variety of long-term fundamental research investments (strategy 1), as well as human-AI collaboration (strategy 2). Several basic research investments are called out as examples in strategy 1: data-focused methodologies for knowledge discovery; perceptual capabilities of AI systems; understanding theoretical capabilities and limitations; general-purpose AI; scalable AI systems; humanlike AI; more capable and reliable robots; hardware for improved AI; and AI for improved hardware.

Separating out the R&D challenges of human-AI collaboration from the other basic research areas was done to place an added emphasis on the need for AI systems to work closely with humans. While there is indeed an important role for fully autonomous AI, such as in self-driving vehicles, many AI systems are designed to augment or enhance the knowledge and capabilities of humans, rather than to replace humans entirely. By creating a strategy specifically focused on AI-human collaboration, the National AI R&D Strategic Plan calls for special attention on cre-

ating AI systems that work effectively in the service of humans. Examples of fundamental AI advances needed in the context of strategy 2 include new algorithms for human-aware AI; AI techniques for human augmentation; visualization and AI-human interfaces; and more effective language processing systems.

Finally, the top row of figure 2 (in darker blue) identifies the broad variety of applications that can benefit from AI advances.

The bottom-up nature of the graphic is intended to convey the fact that a broad approach to foundational AI R&D can be widely beneficial to many applications and that it can be much more effective, as well, than attempting to address research challenges on an application-by-application basis. While no single federal agency will support the entirety of the R&D illustrated in this diagram, the figure does convey a sense of how the federal investments as a whole can work together to address the overall national need. There is not (nor should there be) a one-to-one mapping of research areas to federal agencies, since different mission needs can drive different

areas of emphasis on a particular topic. This graphic also illustrates the potential power of collaboration and coordination of investments across the federal agencies: by working together, the agencies can help ensure that AI R&D is progressing sufficiently in each area to achieve the strategic national goals. This list is not intended to be inclusive.

Measuring Progress, Effectiveness, and Impact of the Strategic Plan

It typically takes years to gauge the impact of a strategic plan in accelerating progress toward defined goals. Key to determining the effectiveness of a strategic plan is measuring progress towards accomplishing the plan's strategic objectives. Given the breadth of topics covered in the National AI R&D Strategic Plan, determining how best to measure progress is difficult. The ongoing development of an AI Index (Shoham 2017) is an important step toward tracking key AI developments in an objective manner. While creating a meaningful AI Index will be a challenge, it could be very helpful in monitoring and measuring progress in AI.

Of course, any plan is only effective to the extent that it is followed and maintained to keep pace with evolving goals and strategic progress. It's important to consider the consequences of increased (or decreased) investments in this area, as well as the balance between government and industry activities. There are many policy makers who believe AI is so important to the nation that it will contribute to long-term economic competitiveness.

It is apparent that many nations of the world have AI near the top of their technology agendas. Most of these decisions are outside of our control as technologists, although perhaps not outside the scope of our influence. In any case, regardless of how the nations of the world follow through on their national AI aspirations, it is incumbent upon us as technologists to focus on the positive, ethical development and use of AI, ensuring that everyone can benefit from the practical application of AI across society, regardless of which nation leads in the strategic development of the technology.

Acknowledgements

Notes

1. The National Artificial Intelligence Research and Development Strategic Plan, October 2016, report from the National Science and Technology Council Networking and Information Technology Research and Development Subcommittee, available at www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf.

2. Preparing for the Future of Artificial Intelligence, National Science and Technology Council report, October 2016, available at obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

References

Felten, E. 2016. Preparing for the Future of Artificial Intelligence. White House Office of Science and Technology Policy (OSTP) blog, May 3, 2016. Washington, DC: Office of the President of the United States. obamawhitehouse.archives.gov/blog/2016/05/03/preparing-future-artificial-intelligence

Felten, E., and Lyons, T. 2016a. The Administration's Report on the Future of Artificial Intelligence. White House OSTP blog, Oct. 12, 2016. Washington, DC: Office of the President of the United States. obamawhitehouse.archives.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence

Felten, E. and Lyons, T. 2016b. Public Input and Next Steps on the Future of Artificial Intelligence. White House OSTP blog, Sept. 16, 2016. Washington, DC: Office of the President of the United States. obamawhitehouse.archives.gov/blog/2016/09/06/public-input-and-next-steps-future-artificial-intelligence

National Research Council. 2012. *Continuing Innovation in Information Technology*. Washington, DC: The National Academies Press. www.nap.edu/catalog.php?record_id=13427

Shoham, Y. 2017. Toward the AI Index. *AI Magazine* 38(4): 71–77. doi.org/10.1609/aimag.v38i4.2761

Stokes, D. E. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, DC: Brookings Institution Press.

Stone, P.; Brooks, R.; Brynjolfsson, E.; Calo, R.; Etzioni, O.; Hager, G.; Hirschberg, J.; Kalyanakrishnan, S.; Kamar, E.; Kraus, S.; Leyton-Brown, K.; Parkes, D.; Press, W.; Saxenian, A.; Shah, J.; Tambe, M.; and Teller, A. 2016. *Artificial Intelligence and Life in 2030*. One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel, Stanford University, Stanford, CA. ai100.stanford.edu/2016-report

Lynne E. Parker is an associate dean of the Tickle College of Engineering at the University of Tennessee, Knoxville, and a professor in the Department of Electrical Engineering and Computer Science. During 2015 and 2016, she served at the National Science Foundation as the division director for information and intelligent systems in the Computer and Information Sciences Directorate. Her service at NSF continued during 2017 as a part-time "expert." During her time at NSF, she served as the coleader of the NITRD Task Force that created the National Artificial Intelligence Research and Development Strategic Plan, and as the NSF representative on the NSTC Subcommittee on Machine Learning and Artificial Intelligence. She received her PhD in computer science from MIT in 1994, and is a fellow of IEEE, a distinguished member of ACM, and a senior member of AAAI.
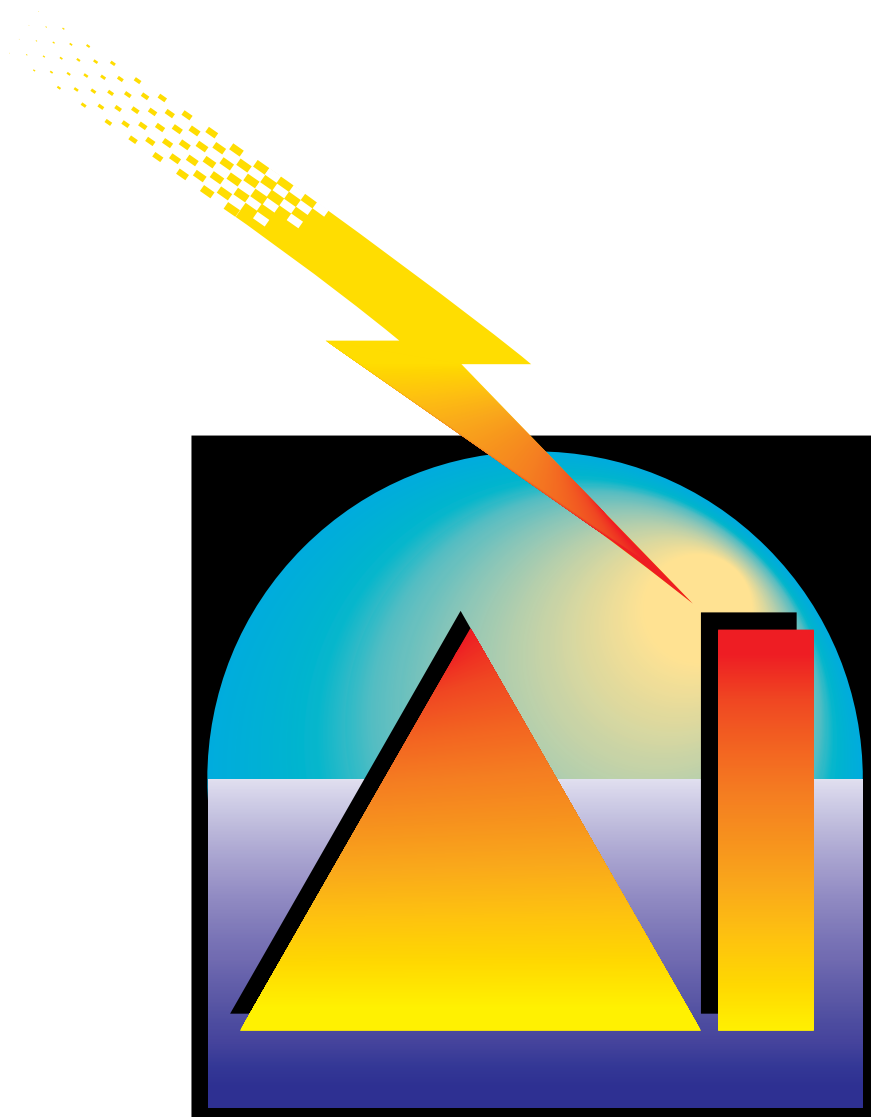
The Thirty-First Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-19)

Honolulu, Hawaii
January 29–31 2019

Karen Meyers, *Conference Chair*
Papers Due September 5, 2018