AI RFI Responses, October 26, 2018

Update to the 2016 National Artificial Intelligence Research and Development Strategic Plan RFI Responses

DISCLAIMER: The RFI public responses received and posted do not represent the views and/or opinions of the U.S. Government, National Science and Technology Council (NSTC) Select Committee on Artificial Intelligence (AI), NSTC Subcommittee on Machine Learning and AI, NSTC Subcommittee on Networking and Information Technology Research and Development (NITRD), NITRD National Coordination Office, and/or any other Federal agencies and/or government entities.
We bear no responsibility for the accuracy, legality or content of all external links included in this document.

©2018 The MITRE Corporation. All Rights Reserved. Approved for Public Release. Distribution Unlimited. Case Number: 18-2319-15

MP180840

October 25, 2018

Response of The MITRE Corporation to the National Science and Technology Council's Request for Information on Artificial Intelligence R&D

For additional information about this response, please contact:
Duane Blackburn, S&T Policy Analyst
The MITRE Corporation
7596 Colshire Drive
McLean, VA 22102-7539

dblackburn@mitre.org
(434) 964-5023

This page intentionally left blank.

Introduction

The MITRE Corporation is a not-for-profit organization that operates federally funded research and development centers (FFRDCs) for the U.S. government. MITRE's FFRDCs serve agencies in a wide range of areas, including national security, education, aviation safety and administration, tax administration, homeland security, health care, benefit services, and other missions. We have directly assisted federal agencies in their strategic investigations of artificial intelligence (AI), and we have also performed independent research to help advance the state of the possible.

AI is going to play a significant role in the Fourth Industrial Revolution[1], where technological capabilities will blur the lines between the physical, digital, and biological spheres. This revolution is happening at a faster pace than the first three and its effects are already being felt in almost every industry. Enhancing AI capabilities and properly applying them (in both government and private sectors) will be critically important for the security and prosperity of the nation.

Both industries and governments around the world, including China and Russia, are significantly investing in AI research. For the U.S. to maintain its technological dominance, it must develop and invest in a sound, integrated, AI R&D strategic plan.

MITRE thus welcomes this opportunity to draw on our technical knowledge and operational experiences to respond to the National Science and Technology Council's (NSTC) Request for Information (RFI) supporting an update of *The National Artificial Intelligence Research and Development Strategic Plan*. MITRE feels that the original (2016) version of this plan was technically accurate and the seven substrategies remain areas that need focused attention. MITRE also believes that the updated strategic plan should continue to follow its predecessor's broader-than-traditional[2] definition of "research and development" to include investigating associated issues that are limiting government's adoption of AI innovations.

MITRE believes, however, that language within the prior strategic plan limited its effectiveness. We are encouraged that the Office of Science and Technology Policy (OSTP) has not only continued the NSTC's work on the subject of AI but also elevated it within the NSTC's structure. We are hopeful that the updated strategic plan follows suit, becomes influential, and provides several ideas within this response to enhance its effectiveness.

Recommendation #1: Better Leverage the Power of the NSTC

The 2016 version of the AI R&D Strategic Plan includes statements that limited its effectiveness. For example, "The AI R&D Strategic Plan does not, however, define specific research agendas for individual Federal agencies. Instead, agencies will continue to pursue priorities consistent with their missions, capabilities, authorities, and budgets, while coordinating so that the overall research portfolio is consistent with the AI R&D Strategic Plan." Statements such as these undercut the

[1] Schwab, Klaus."The Fourth Industrial Revolution: what it means, how to respond," World Economic Forum, January 2016 (https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/ accessed October 16, 2018)
[2] Basic Research, Applied Research & Development

© 2018 The MITRE Corporation.
All rights reserved.

ability of the NSTC to drive research and ensure that it meets strategic aims[3], and replaces it with a follow, analyze, and report function.

Executive Order 12881 states, in part, that the functions of the NSTC are to "ensure science and technology policy decisions and programs are consistent with the President's stated goals" and to "provide advice to the Director of the Office of Management and Budget (OMB) concerning the agencies' research and development budget submissions." Over the past 25 years, the NSTC's biggest impacts have occurred when it develops solid whole-of-government R&D plans, which are then reinforced by OSTP and OMB staffers during budget development and execution. Given the importance of AI to the nation's future and the breadth of activities taking place throughout the Executive Branch, MITRE recommends that the updated AI R&D strategic plan better leverage these NSTC capabilities by:

- Updating the existing strategy areas so they are no longer presented as ideals for agencies to consider but instead as specific outcomes that the government intends to meet;
- Assigning a lead agency, as well as supporting agencies, for each outcome;
- Tasking the Select Committee on AI to regularly track progress on each outcome, and report on such to OSTP and OMB;
- Directing agencies to support the strategic plan within their R&D budgets; and
- Stating an intent for OSTP and OMB to closely analyze agencies' budget requests to ensure proper support for this strategic endeavor.

There is ample precedent of NSTC strategies and subcommittees performing such activities. Given the NSTC's recent elevation of its AI activities, it makes sense for this subject to now receive commensurate levels of attention and support.

Recommendation #2: Seek Collaboration with Non-Government Entities on AI R&D

The 2016 AI R&D strategy states that the U.S. government's investment in unclassified AI R&D during 2015 was approximately $1.1 billion, and a May 2018 Fact Sheet[4] states that the government's investment has grown by over 40 percent since that time—equating to just over $1.5 billion. In contrast, U.S. venture capital investments in AI for 2017 is estimated at $6 billion[5], which is approximately four times the government's investment. Large companies' internal investments[6] aren't included in this number, which makes the government's percentage of the overall national contribution even lower. To be a true *national* R&D plan and, more importantly, to better meet the nation's strategic goals for AI, non-government investments and activities must also be integrated into the strategy. To support the role of non-government entities in this endeavor, the U.S. government can develop a holistic strategy, create and/or participate in public-private

[3] See https://www.mitre.org/sites/default/files/publications/pr-16-0916-interagency-s-and-t-leadership.pdf for a discussion on how NSTC activities can be led to impact federal agency budgets and activities.
[4] https://www.whitehouse.gov/briefings-statements/artificial-intelligence-american-people/ (accessed October 8, 2018)
[5] https://www.krostcpas.com/ma-news/global-investment-trends-artificial-intelligence-machine-learning (accessed October 8, 2018)
[6] Apple's investment is estimated at approx. 1/3 that of the U.S. Government (https://www.marketwatch.com/story/ai-in-2018-apple-is-acquiring-knowledge-but-results-may-take-a-while-2017-12-28, accessed October 16, 2018); Microsoft's Artificial Intelligence & Research Group is staffed with 5,000 scientists & engineers (https://www.fool.com/investing/2018/06/20/heres-how-microsoft-is-investing-in-ai.aspx, accessed October 16, 2018), whose salary total would roughly equal the U.S. Government's stated investments.

© 2018 The MITRE Corporation.
All rights reserved.

partnerships[7], track non-government investments, and devote its resources to overcoming priority gaps in the nation's investments.

A challenge to implementing this recommendation will be finding ways to incentivize non-government entities to collaborate. The federal government is neither the intellectual leader nor predominant sponsor in AI R&D, so the incentives that the government normally relies on aren't relevant here. The problem is compounded within this specific topic as many community leaders are hesitant to see AI used within government contexts. The government will need to address this concern within its strategy and implementation plan.

Recommendation #3: Expand Research into Cognitive, Psychological, and Social Effects of Artificial Intelligence

The NSTC's 2016 Strategic Plan emphasizes the need to understand and address the ethical, legal, and societal implications of AI (strategy 3). This plan includes three areas for focus: 1) Improving fairness, transparency, and accountability-by-design, 2) Building ethical AI, and 3) Designing architectures for ethical AI. MITRE believes that these areas remain critical aspects of a national AI R&D strategy and recommends adding another research area: the social effects (including related cognitive and psychological effects) of introducing AI agents into existing governmental organizations and other social structures. While technology frequently impacts social structures, the introduction of AI agents raises particular opportunities and issues that must be understood more fully, both for policymaking and to guide the development of government systems. These issues will become more pressing as AI systems increase in capability and perceived intelligence.

In some ways, the social effects of AI systems echo of social impact of introducing computer systems in past decades. The introduction of computers, and their gradual increase in ubiquity, has changed how we do business, handle governmental deliberations, and interact with one another. Similarly, current AI systems—including chatbots, voice assistants, decision support systems, and automated legal and tax assistants—have changed and will continue to change government's social interactions. This will be particularly true as they provide capabilities that more closely replicate the abilities of humans. As these new capabilities allow AI systems to handle tasks of much greater complexity and contextual variability, human users may treat the systems as if they are persons, even while understanding they are not. Computer systems have been designed and perceived as tools, but many AI systems already have been designed to reflect human-like interaction patterns, which can blur the person/tool distinction.[8]

This blurring between AI as tool and AI as "person" must be better understood as it affects the usefulness of AI systems. As an example, consider bias in machine learning results for predicting recidivism in convicts. Most machine learning systems derive much of their flexibility and value from extensive training data. This training data provides for a system that can respond appropriately in many disparate contexts, as a human judge might, but also raises the issue of potential bias in some cases. Some machine learning-based recidivism models have been shown to

[7] The Computing Community Consortium, for example, has already stated a desire to bring together academic and industrial researchers with federal agency representatives to establish a "Roadmap for Artificial Intelligence." See http://www.cccblog.org/2018/09/20/a-roadmap-for-artificial-intelligence/ (accessed October 18, 2018).
[8] Cofone, I. N. "Servers and Waiters: What Matters in the Law of A.I." *Stanford Technology Law Review*. 21(2). September, 2018. https://law.stanford.edu/publications/servers-and-waiters-what-matters-in-the-law-of-a-i/. (accessed October 16, 2018)

© 2018 The MITRE Corporation.
All rights reserved.

exhibit racial bias[9], and this problem has recurred in multiple contexts.[10] This bias is related to human bias (as reflected in the training data itself), but also goes beyond that bias. To treat machine learning-based bias as equivalent to human bias is tempting. However, that reflects an incomplete understanding of the causes, which may also be influenced by other factors, such as undersampling of minority groups.[11] Thus to understand this issue requires not treating the bias as a result of human bias or as a broken mathematical tool, but as something in the middle that requires multiple fixes.

Additional examples include:

- The use of AI to act as agents of deception in social media. Recent research has shown that AI-based bots can influence the direction of conversations on topics in social media,[12] and that some foreign governments have employed them to do so.[13] It is important to note that such bots do not need to be wholly convincing replacements for humans; they may be successful if they merely disrupt or cast doubt about the validity of participants.
- The effects of interactions with AI agents on social isolation, either positive or negative. Research is being conducted on adapting social agents to help detect depression in humans through machine learning-driven analyses of discussions.[14]
- Government or other organization AI agents (such as chatbots or voice assistant skills). There is a need to understand what AI-linked characteristics make an organization feel responsive (e.g., by making available a variety of services 24x7) and which create a sense of disconnection with the organization (by placing an additional layer between the human user and a human organization representative).
- The side effects of machine-learning-driven recommendation systems that attempt to maximize engagement. For example, it has been suggested that YouTube's recommendation algorithm[15] may lead users to gradually view videos with more and more extreme views in an effort to maximize viewing time.[16] For situations in which that outcome is suboptimal from a societal point of view, what are effective countermeasures?
- Exploring how to best transition into human-machine teaming situations so that people don't feel AI is a threat to their jobs, but rather is an enabler that will help them perform better.

---

[9] Larson, Jeff, Mattus, Surya, Kirchner, Lauren and Angwin, Julia. "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, May 23, 2016. https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. (accessed October 16, 2018)

[10] Cossins, Daniel. "Discriminating algorithms: 5 times AI showed prejudice." *New Scientist*. April 27, 2018. https://www.newscientist.com/article/2166207-discriminating-algorithms-5-times-ai-showed-prejudice/. (accessed October 16, 2018)

[11] Buolamwini, J., & Gebru, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification.* Paper presented at the Conference on Fairness, Accountability and Transparency.

[12] Watts, Clint. "Artificial intelligence is transforming social media. Can American democracy survive?" *The Washington Post*. September 5, 2018. https://www.washingtonpost.com/news/democracy-post/wp/2018/09/05/artificial-intelligence-is-transforming-social-media-can-american-democracy-survive. (accessed October 16, 2018)

[13] Dougherty, Jill, and Jay, Molly. "Russia tries to get smart about artificial intelligence." *The Wilson Quarterly*. Spring, 2018. https://wilsonquarterly.com/quarterly/living-with-artificial-intelligence/russia-tries-to-get-smart-about-artificial-intelligence/. (accessed October 16, 2018)

[14] Robinson, Ann. "Meet Ellie, the machine that can detect depression." *The Guardian*. September 15, 2017. https://www.theguardian.com/sustainable-business/2015/sep/17/ellie-machine-that-can-detect-depression. (accessed October 16, 2018)

[15] Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., . . . Livingston, B. (2010). *The YouTube video recommendation system.* Proceedings of the Fourth ACM Conference on Recommender Systems.

[16] Nicas, Jack (2018-02-07). How YouTube Drives People to the Internet's Darkest Corners. T*he Wall Street Journal. url: https://www.wsj.com/articles/how-youtube-drives-viewers-to-the-internets-darkest-corners-1518020478.* (accessed October 16, 2018)

© 2018 The MITRE Corporation.
All rights reserved.

For these reasons, we suggest that the federal government should expand research into the cognitive, psychological, and societal effects of AI use.

Researching these issues at the federal level is important because many of these problems relate to citizens' quality of life. The private sector may not address quality of life concerns if they don't directly affect the profitability of AI-based systems. There is little incentive, for example, for recommendation systems to provide results that reduce engagement rather than increase it. By expanding research into the social effects of AI use, policymakers and system designers in the federal government will be able to manage these emerging issues in an effective way.

Recommendation #4: Adapting Infrastructure to Provide Affordances for AI Systems

The current *National Artificial Intelligence R&D Strategic Plan* touches on AI research infrastructure within a few of its strategies, including ensuring safety and security (Strategy 4), development of shared datasets and environments (Strategy 5), and AI standards and benchmarks (Strategy 6). Falling between the cracks, however, is an important infrastructure topic that could be aided by federal support: the inclusion of artificial intelligence affordances in the built infrastructure could enhance the use of AI products.

Affordances are features of objects that enable people to use them. Good affordances make it easier for people to use objects in the way they want or need. Examples include handles on frying pans and doorknobs on doors; traffic lights and exit signs are affordances that enable motorists to use the roadways. As we introduce AI into products and systems, in addition to making them easy to use, we also want them to operate reliably. To enable this, we may need to modify our environments.

Consider transportation as an example. The Vision mentions: "structural health monitoring," "increasing situational awareness," and "provide drivers and other travelers with real-time route information." These actions all involve, or would be aided by, building sensors and signals that automated systems will be able to easily interact with in our infrastructure. If "situational awareness" and "real-time route information" include the state of traffic signals, construction zones, and road conditions, then we should be looking for ways to make it simple for automated systems to query those things. Some observers consider China to be very effective at this.[17]

The potential benefits of such AI-friendly affordances can be observed in robot competitions, such as robot soccer.[18] There are many levels of competition ("leagues"), each with its own playing fields. For example, in the small size competitions, the field has specified dimensions and color (green with white lines) with goals of given dimension and location. The field of play is surrounded by a low wall. The ball is an orange golf ball. Additionally, a shared, open-source vision system, maintained by the playing community, is keyed to these preset environmental characteristics. The system uses an overhead camera that provides location information to each team, including the location of the ball and locations of all players on the field. Players are required to carry color markings of specified size and color. These conditions allow the player logic to more easily find

[17] K-F Lee, What China Can Teach the U.S. About Artificial Intelligence, Opinion, *The New York Times*, September 23, 2018.
[18] https://www.robocup.org/. (accessed 12 October 2018.)

© 2018 The MITRE Corporation.
All rights reserved.

other robots, the ball, the goal, and the field boundaries. Some of them are very similar to what is found in human soccer (e.g., white lines, white goal post, distinct player uniforms) but some are very different (identical field sizes in every event, orange ball, shared overhead vision).

These conditions allow the developers to focus on intelligent multi-agent cooperation in a dynamic environment and expend less effort on vision processing and boundary detection. These latter issues, which might be useful to solve in more general ways, are relatively easy to simplify out of the robot soccer environment.

Autonomous automobile operation on our common roads is analogous to RoboCup. It should be possible to provide recognizable clues in the road environment for things such as lane boundaries and traffic signals, and maybe even other cars. But the dynamics of other drivers, pedestrians, and changing light and weather conditions are much harder to handle. By making the simple aspects easy, we would enable the AI designers and developers to focus attention on the complex ones.

MITRE recommends expanding Strategy 6 to investigate the benefits and best practices of incorporating affordances within AI systems.

Research questions:

- How can our physical infrastructure be modified to make it easier for AI to operate with it?
- How can this be done without interfering with the needs of humans to use the infrastructure?
- How can we upgrade our infrastructure to not only better enable AI integration but also to protect it from misfunctioning AI implementations?

Outcome: Safer and more intelligent AI will arrive faster when more of the operating environment is easier for systems to handle.

Recommendation #5: Investigate Implementation Issues that Are Slowing the Impact of AI R&D Gains.

Advances from prior AI R&D have far outpaced government adoption of AI capabilities.[19] Increasing attention and funding for AI R&D, without addressing the lab-to-field disconnect, will enlarge this chasm, thus decreasing the return on the government's R&D investments. As part of this strategic plan, the NSTC, working with its peers whenever possible, should also seek to understand and overcome these implementation issues so that tomorrow's innovations can be rapidly transitioned to serve the nation.

Barriers and areas of concern include:

- Understanding AI performance. Today, the ability to analyze the decisions of AI systems and measure their accuracy and reproducibility is limited. The combination of the federal

[19] https://about.bgov.com/blog/pentagon-bridge-artificial-intelligences/. (accessed October 16, 2018)

© 2018 The MITRE Corporation.
All rights reserved.

bureaucracy's risk-averse culture[20] with the public's distrust of AI[21] (particularly the federal government's use of AI[22]) serve as significant limiters to federal AI adoption.

- Immature guidance on legal and ethical considerations. "Regulatory and social barriers can raise the cost and slow the rate of adoption. Product liability is a concern. Privacy considerations restrict access to data and often require it to be anonymized before it can be used in research. Ethical issues, such as trained biases and algorithmic transparency, also remain unresolved."[23]
- Data sharing limitations. Even the most-advanced AI systems in the world will be virtually useless without access to necessary operational data, which is often unavailable due to statutory limits, privacy concerns, or other equally relevant considerations. The Select Committee should partner with leadership of the President's Management Agenda CAP Goal #2[24] to ensure proper access and adequate safeguards are in place for the use of future AI systems.
- There is a large, and growing, AI skills gap[25]. All entities "are struggling to secure employees with the combination of skills and knowledge necessary to unleash the full potential of AI."[26] The nation isn't producing enough new graduates with these skills,[27] and most of the graduates we are producing are entering the private sector or going to work for our nation-state competitors.
- The next generation will ultimately lead us. In the 1990s, students' exposure to HTML and web technologies led to a transformational shift in the way companies and the government did business in the 2000s. The same will likely hold true for AI. We need to ensure our next-generation workforce has an early exposure to AI across a wide variety of fields of study.

---

[20] https://www.mitre.org/sites/default/files/publications/pr-15-3060-innovation-landscape-government-future-role_0.pdf

[21] https://www.scientificamerican.com/article/people-dont-trust-ai-heres-how-we-can-change-that/. (accessed October 16, 2018)

[22] "In the end, public trust will be maintained through demonstrating that the technology is beneficial and that safeguards work. This will require, at a minimum:

- Correctly identifying any harmful impacts of artificial intelligence.
- Formal structures and processes that enable citizen recourse to function as intended.
- Appropriate means of redress.
- Clear accountability.
- Clearly communicating the substantial benefits for society offered by artificial intelligence."

Artificial intelligence: opportunities and implications for the future of decision making, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf

[23] https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx. (accessed October 16, 2018)

[24] https://www.performance.gov/CAP/CAP_goal_2.html. (accessed October 16, 2018)

[25] https://www.forbes.com/sites/bernardmarr/2018/06/25/the-ai-skills-crisis-and-how-to-close-the-gap/#3bb41b8831f3. (accessed October 16, 2018)

[26] https://www.rdmag.com/article/2018/05/four-key-barriers-widespread-adoption-ai. (accessed October 16, 2018)

[27] https://www.mitre.org/publications/project-stories/preparing-generation-ai-for-the-fourth-industrial-revolution. (accessed October 16, 2018)

© 2018 The MITRE Corporation.
All rights reserved.