

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.



October 26, 2018

Faisal D'Souza
Technical Coordinator, Artificial Intelligence R&D Interagency Working Group (IWG)
NCO for NITRD
2415 Eisenhower Avenue
Alexandria, VA 22314

RE: Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO), National Science Foundation Notice of Request for Information: Request for Information on Update to the 2016 National Artificial Intelligence Research and Development Strategic Plan (FR Doc. 2018–20914)

Dear Mr. D'Souza:

On behalf of the American Psychological Association (APA) and its 115,700 members and affiliates, we would like to thank the National Science Foundation (NSF) and the Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO) for the opportunity to provide input on the updated National Artificial Intelligence Research and Development Strategic Plan. The National Artificial Intelligence Research and Development Strategic Plan works to help coordinate the nation's efforts on artificial intelligence (AI) along seven strategic aims. As noted in the RFI, AI has the potential to rapidly transform numerous aspects of society, making thoughtful priorities and progress essential for guaranteeing pro-social outcomes and meaningful advances for people across the nation.

The American Psychological Association is the largest scientific and professional organization representing psychology in the United States, with a mission to advance the creation, communication, and application of psychological knowledge to benefit society and improve people's lives. In that spirit, we submit the following comments and eagerly await additional detail on the updated National Artificial Intelligence Research and Development Strategic Plan. We look forward to maintaining a dialogue with the White House Office of Science and Technology Policy (OSTP), the National Science and Technology Council (NSTC), NITRD, and NSF as investment and development of AI continue with the goal of increasing its social benefit and effectiveness through the application of psychological science.

Beginning on the following page, we are pleased to offer comments on the strategic plan priorities.

750 First St. NE
Washington, DC 20002-4242
(202) 336-6000
www.apa.org

Strategy 1: Make long-term investments in AI research.

- Psychologists have been engaged in work on AI, or the core principles that inform AI, for decades, dating back as far as the 1950s. For example, psychologist Margaret Boden’s work on AI throughout the 1980s helped frame the contributions of psychological science alongside the work of Turing, von Neumann and other pioneers in automation and AI. Psychologists have had tremendous influence on AI through both bottom-up “connectionist” research, by B.F. Skinner and others, as well as research on top-down processing (Gopkin, 2017).
- Achievement of the long-term priorities from the existing strategic plan require current and future continued investments in psychological science across the research spectrum, ranging from basic research to applied work. The current strategic plan’s stated priorities of enhancing the perceptual capabilities of AI systems, pursuing research on general-purpose artificial intelligence, and fostering research on human-like AI all mandate the inclusion of psychological science.
- A breadth of psychology research has supported AI from its inception. This work includes contributions for all disciplines of psychology, including cognitive, personality, social, neuropsychology, developmental, clinical, and experimental psychology. Each discipline of the field provides a unique perspective and contribution to the ongoing development of AI.
- Importantly, as the current strategic plan notes, there is often a “long, unpredictable incubation period” between original research and its eventual application. Failure to properly support basic and applied psychological research today has the potential to slow, and even entirely prevent, the AI breakthroughs of tomorrow.

Strategy 2: Develop effective methods for human-AI collaboration.

- As noted in the strategic plan, “While completely autonomous AI systems will be important in some application domains, many other application areas are most effectively addressed by a combination of humans and AI systems working together to achieve application goals.” Human/AI cooperation is one area in which psychological science has a unique and vital contribution.
- Theory of mind, understanding the internal states and beliefs of others, is essential for AI to function alongside humans in a productive capacity. Although even young children possess this ability, AI can struggle to perform tasks requiring a working theory of mind. Work by a number of teams of psychologists and computer scientists (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Rabinowitz et al., 2018; Raileanu, Denton, Szlam, & Fergus, 2018) has demonstrated inroads on helping AI more accurately infer internal states of other actors and correctly predict future behavior by working to replicate theory of mind.

- Psychologists also have experience working alongside AI systems, complementing the expertise of each. Walsh, Ribeiro, & Franklin (2017) combined clinical psychological practice and machine learning to predict suicide attempts among individuals at risk for self-injurious behavior. Clinical psychological practice is ripe for AI-human collaboration with opportunities in “clinical training, treatment, psychological assessment, and clinical decision making,” however, trained psychologists are essential for consideration of individual differences, cultural contexts, and therapeutic relationships (Luxton, 2013).
- Human behaviors as subtle as eye contact are vital to cognitive processes and action. Psychologists work to examine how these behaviors apply in the context of human-AI interaction, such as autonomous vehicles (Gredebäck & Falck-Ytter, 2015; Vernon, von Hofsten, & Fadiga, 2011) and social robotics (DeSteno, Breazeal, Frank, Pizarro, Baumann, Dickens & Lee, 2012).

Strategy 3: Understand and address the ethical, legal, and societal implications of AI.

- Similar to other algorithm-based systems, AI is vulnerable to the biases and weaknesses of the underlying data. Numerous high-profile examples, such as Microsoft’s chat bot, Tay, have demonstrated shortcomings in the algorithmic processing AI rely upon.
- AI, and machine learning, can quickly adopt the implicit biases existing within seemingly innocuous language (Caliskan, Bryson, & Narayanan, 2017). Importantly, work demonstrating this vulnerability is based on the pioneering research of psychologists attempting to measuring individual differences in automatic cognition (Greenwald, McGhee, & Schwartz, 1998).
- Psychologists are the leading authority on understanding the mechanisms underlying bias and, ultimately, strategies attempting to reduce bias and promote equity. Any ethical implementation of AI must address these biases to ensure that AI utilization does not exacerbate existing social and structural inequality and inequity.

Strategy 4: Ensure the safety and security of AI systems.

- Understanding the inner mechanisms of AI functioning will require the participation of psychologists. The algorithmic processing of neural networks and AI provide little external output and are often considered “black boxes.” Researchers are more commonly adapting cognitive tests and measures for assessing AI processes. DeepMind’s PsychLab is an open-source testing environment used to enable traditional psychological testing for AI and other autonomous systems (Leibo et al., 2018). Other researchers have used psychological research on learning and memory as a base for testing how well AI generalizes performance across variable novel environments (Oh, Chockalingam, Singh, & Lee, 2016).

- Researchers have posited that AI could exhibit maladaptive cognitions, akin to pathological conditions, and argue for a psychopathological approach to deleterious AI behaviors (Behzadan, Munir, Yampolskiy, 2018). Other researchers have suggested sufficiently advanced AI may also be vulnerable to conditions analogous to depression (Mainen, 2018). A team at the Massachusetts Institute of Technology (MIT) trained an AI they dubbed the “first psychopathic AI” using images from graphic online communities. As AI becomes more prevalent and complex, the pressing need for psychological expertise will continue to accelerate.

Strategy 7: Better understand the national AI R&D workforce needs.

- Psychologists are an essential component of both the current and future AI R&D workforce. Unfortunately, psychologists are often given insufficient attention for their contributions in AI-related work, despite the fundamental work on AI being psychological in nature.
- As AI continues to develop and become more complex, psychologists will continue to become more essential, particularly in regard to discerning the internal mechanisms of algorithmic processing and promoting human-AI collaboration. Current work at NSF, Growing Convergence Research under the 10 Big Ideas umbrella, focuses on fostering research moving beyond simply interdisciplinary efforts to infuse transdisciplinary thought and theory into each aspect of the research paradigm. Similar efforts should be applied to future workforce training to ensure that psychologists, computer scientists, and other AI stakeholders receive sufficient training and experience to allow for convergent AI research that meets the needs of the nation.

We hope this information is helpful. If there are additional areas in which APA can assist or provide support in the ongoing update of the National Artificial Intelligence Research and Development Strategic Plan, please contact Dr. Steve Newell, Senior Legislative and Federal Affairs Officer, APA Science Directorate, at Snewell@apa.org or (202) 507-7175.

Sincerely,

Howard S. Kurtzman, Ph.D.
Acting Executive Director for Science