

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.

NISKANEN C E N T E R

Regulatory Comment

Comments submitted to the National Science Foundation in the Matter of:

REQUEST FOR INFORMATION RESPONSE: NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN

Ryan Hagemann
Senior Director for Policy
Niskanen Center

Submitted: October 23, 2018
Docket Number: 2018-20914

EXECUTIVE SUMMARY

The 2016 National Artificial Intelligence Research and Development Strategic Plan was a commendable step forward in crafting a national strategy for research and development in artificial intelligence. Given the intense investments being made by China and other countries, it is imperative that the United States not only matches, but surpasses these efforts. To achieve a vision of continued American leadership in technological progress and innovation, the National AI Strategy requires some updates to maximize the effectiveness of these investments. These comments will argue in favor of two such updates.

First, the National AI Strategy should include an additional strategic aim that focuses on how best to remove burdensome regulatory barriers, which present unnecessary obstacles to ongoing research, testing, and commercial deployment of new artificial intelligence systems and technologies. Second, the third strategy should be expanded to include a consideration of the trade-offs associated with incorporating ethical frameworks into the architectural design of artificial intelligence systems. In particular, we recommend explicitly broadening the strategic aim to consider outcomes-based governance strategies that correct specific, identifiable harms that may result from the application of autonomous decision-making systems.

INTRODUCTION

In 1989, engineers at Carnegie Mellon University made history when they successfully navigated the first semi-autonomous vehicle — specifically, an Autonomous Land Vehicle In a Neural Network, or ALVINN — along a roadway, using nothing but “images from a camera and a laser range finder.”¹ The researchers concluded that these findings could serve as a proof-of-concept for “the possibility of a novel adaptive autonomous navigation system capable of tailoring its processing to the conditions at hand.”² In the early 2000s, building on these decade-old experiments, the Defense Advanced Research Projects Agency (DARPA) conducted a series of Grand Challenges — long-distance competitions aimed at assessing the technical viability of autonomous transportation technology, as well as spurring further developments in the robotics and sensor technologies that would be required for true autonomous vehicle capabilities.³

Today, autonomous vehicles have leaped from the pages of science fiction novels into everyday headline news. While this technology holds the potential to completely reimagine the future of American transportation networks, its adoption is anything but assured. Even as the DARPA Grand Challenges helped accelerate many of the technical developments that have made the coming age of autonomous transportation possible, the legal and regulatory challenges to those deployments continue to plague developers and policymakers alike. Had more policy consideration been given to the advent of autonomous vehicles earlier in the technology’s life cycle, these non-technical challenges could have potentially been ameliorated or significantly defrayed. Now, as we look towards a future of applied artificial intelligence (AI) beyond the transportation sector, the government has a second chance at considering the policy implications of this new technology. It should not repeat the same mistakes made in the early years of autonomous vehicles.

To that end, the 2016 National Artificial Intelligence Research and Development Strategic Plan (hereafter, “National AI Strategy”) outlined seven general strategies for how the government could help yield research that would “produce new AI knowledge and technologies that provide a range of positive benefits to society, while minimizing the negative impacts.”⁴ Those strategies included:

1. Make long-term investments in AI research.
2. Develop effective methods for human-AI collaboration.
3. Understand and address the ethical, legal, and societal implications of AI.
4. Ensure the safety and security of AI systems.
5. Develop shared public datasets and environments for AI training and testing.
6. Measure and evaluate AI technologies through standards and benchmarks.
7. Better understand the national AI R&D workforce needs.

¹ Tom Vanderbilt, “Autonomous Cars Through the Ages,” *Wired*, 6 Feb. 2012, <http://www.wired.com/autopia/2012/02/autonomous-vehicle-history>.

² *Id.*

³ Cyber Grand Challenge, DARPA, <http://www.cybergrandchallenge.com>.

⁴ *The National Artificial Intelligence Research and Development Strategic Plan*, National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, Oct. 2016, p. 3 (hereafter, *National AI Strategy*), https://www.nitrd.gov/pubs/national_ai_rd_strategic_plan.pdf.

Although these strategies are largely commendable, there are some gaps that will need to be plugged to minimize the potential downsides while maximizing the society-wide benefits. These comments will focus on the need for expanding and updating the research agenda detailed in Strategy 3, as well as adding a strategy to address the legal and regulatory barriers to AI research and development, commercialization, and adoption.

PART I: ADDRESSING BARRIERS TO INNOVATION

The National AI Strategy explicitly narrowed its mission to focus “on the R&D investments needed to help define and advance policies that ensure the responsible, safe, and beneficial use of AI.”⁵ However, a discussion of the many applicable uses of AI spurred by federal R&D investment is incomplete without addressing the potential legal and regulatory challenges likely to be faced by more widespread and commercial deployment of AI systems. Fortunately, the Office of Science and Technology Policy (OSTP), which jointly authored the National AI Strategy alongside the National Science and Technology Council (NSTC) Subcommittee on Networking and Information Technology Research and Development (NITRD), is well-positioned to address those barriers to technological innovation.

Under its originating statute, OSTP is charged with “provid[ing] the President with periodic reviews of Federal Statutes and administrative regulations of the various departments and agencies which affect research and development activities, both internally and in relation to the private sector, or which may interfere with desirable technological innovation, together with recommendations for their elimination, reform, or updating as appropriate.”⁶ This provision is no accident; indeed it was echoed as a priority for OSTP’s predecessor, the President’s Committee on Science and Technology.⁷

Further, as noted by Executive Order No. 12881, the NSTC is tasked with “coordinat[ing] the science and technology policy-making process.”⁸ As such, NSTC is a suitable forum for promoting OSTP’s statutory missions to provide recommendations for the “elimination, reform, or updating” of “administrative regulations of the various departments and agencies which affect research and development activities ... which may interfere with desirable technological innovation.”⁹ Thus, it is appropriate for the National AI Strategy to include an additional strategic priority that focuses on identifying, reviewing, and suggesting changes to the U.S. Code of Federal Regulations (CFR) that may impede the development of AI in both government and the private sector.

PART II: RECOMMENDATIONS

Add an Eighth Strategic Aim to Address the Legal and Regulatory Barriers to the Research, Deployment, and Adoption of AI

In the interest of ensuring the United States continues to lead the world in technological innovation, the National AI Strategy should be updated to include an additional strategic aim that focuses on the need for

⁵ *Id.*, p. 8.

⁶ 42 U.S.C § 6614(a)(7).

⁷ Prior to the transfer of its authorities to OSTP, the President’s Committee on Science and Technology was tasked with the “reduction and simplification of Federal regulations and administrative practices and procedures which may have the effect of retarding technological innovation or opportunities for its utilization.” 42 U.S.C § 6633(a)(2)(4).

⁸ E.O. 12881, § 4(a)(1).

⁹ 42 U.S.C § 6614(a)(7).

removing regulatory barriers that might present obstacles to researching, testing, and ultimately deploying new AI systems. This strategy could read as follows:

Strategy 8: Review and identify administrative rules, regulations, and policies that may hinder the research, deployment, and adoption of AI systems within the federal government and private sector.

For all the reasons outlined in Part I, the addition of this strategic aim is not only appropriate given the statutory mission of OSTP, but also necessary to ensure the United States does not fall behind in the global AI race. It is also a necessary precondition to ensure the government is capable of meeting all of the other seven strategic aims outlined in the National AI Strategy. After all, the ultimate value of AI is in its actual implementation, not in merely acquiring an understanding of its implications. Unless the government prioritizes the identification and reformation of rules and regulations standing in the way of commercial AI deployment, the optimal benefits promised by this technology will continue to elude society.

Expand the Scope of Strategy 3 to Address the Merits of “Algorithmic Accountability” via a Multistakeholder Process

Strategy 3, as detailed in the 2016 Strategic Plan, focuses on the need for research that can help “[u]nderstand and address the ethical, legal, and societal implications of AI.” In particular, it notes:

A multi-disciplinary approach is needed to generate datasets for training that reflect an appropriate value system, including examples that indicate preferred behavior when presented with difficult moral issues or with conflicting values. These examples can include legal or ethical “corner cases”, labeled by an outcome or judgment that is transparent to the user. AI needs adequate methods for values-based conflict resolution, where the system incorporates principles that can address the realities of complex situations where strict rules are impracticable.¹⁰

While considering the viability of architectural designs for ethical AI is certainly a noble goal, such an agenda presupposes both the technical feasibility and inherent value of such designs. Rather than focusing narrowly on “values-based conflict resolution” systems, Strategy 3 should be broadened to consider ex post facto governance strategies for addressing known, quantifiable, and clearly-identifiable harms resulting from the use of AI systems. As detailed in a forthcoming white paper from the Niskanen Center, the principles of *algorithmic accountability* could potentially address the many “ethical, legal, and societal implications of AI,” without the need for ex ante architectural design frameworks:

Some have argued the best way of achieving these ends is through “algorithmic transparency” — a term lacking in precise meaning, but which generally holds the need to peer through the “black box” of neural networks in an attempt to understand why particular automated decisions might have been made. Although well-intentioned, such an approach is all but impossible given current technologies; and even if it were possible, rules that subject the inner workings of an AI system and its code to pre-review/pre-approval would almost certainly disincentivize research and development into proprietary AI technology. Luckily, there is a more practical and effective regulatory alternative that can protect consumers while continuing to promote innovation and research.

Instead of requiring “algorithmic transparency,” policymakers should insist on “algorithmic accountability.” In practice, this would mean that the operator of an AI decision-making algorithm would be responsible for the consequences of that decision. The operator would be expected to provide

¹⁰ National AI Strategy, p. 27.

*a non-technical, high-level explanation for a particular decision upon request, which might include a confidence interval and a demonstration of procedural regularity.*¹¹

Currently, the scope of Strategy 3 aims to address not only how to design “justice and fairness” into AI systems — “within the bounds of current engineering techniques”¹² — but also implicitly demands that researchers solve age-old epistemic and metaphysical questions about “justice and fairness.” These are nearly insurmountable tasks that will invariably prioritize certain values over others, potentially to the detriment of large swaths of the American population. A better approach would be to expand Strategy 3 to examine how ethically-neutral design standards would compare to a framework that prioritizes a static set of ethical value judgments.

Insofar as the National AI Strategy recognizes that “[r]esearch in this area can benefit from multidisciplinary perspectives that involve experts from” a wide array of fields, Strategy 3 should also be expanded to explicitly call for these voices to be heard through a multistakeholder process, potentially convened by OSTP, NSF, or both.¹³ This reorientation will ensure a wider range of views and potential solutions are given greater consideration in research funding and regulatory coordination strategies.

CONCLUSION

The DARPA Grand Challenges helped spur the research and innovation that helped catalyze the emergence of autonomous vehicle research and testing. Unfortunately, the lack of an early concerted effort focused on identifying and updating existing barriers to that technology’s emergence has put the Department of Transportation in the unenviable position of attempting to craft rules in the midst of a rapidly changing technological landscape. Regulators continue to be caught flat-footed by the emergence of new technologies that evolve at such a speed that the law and regulations cannot keep pace. Had earlier consideration been given to the need to reform the Federal Motor Vehicle Safety Standards and other related portions of the CFR, autonomous vehicle technology might have had a quicker path to commercialization.

The NITRD, NSTC, and OSTP would do well to heed the lessons of the government’s initial investment in autonomous vehicle technology. In updating the National AI Strategy, the government has an opportunity to get out ahead of these potential regulatory roadblocks. While it is not possible to accurately diagnose all potential future problems that may result from the widespread use of AI, preemptively addressing the rules that may hinder this technology’s deployment is well within the ability of the government. OSTP must help lead this effort, as detailed under its originating statute, by pinpointing those areas of the CFR and other regulations that should be updated to help realize the benefits AI can offer the American people.

We would like to thank NSF for the opportunity to comment on this issue and look forward to continued engagement on this and other topics.

¹¹ Ryan Hagemann, et. al., “The Policymaker’s Guide to Emerging Technologies,” Niskanen Center (Washington, DC.: Nov. 2018).

¹² *National AI Strategy*, p. 26.

¹³ *Id.*, p. 26. For a more thorough accounting of the costs and benefits of multistakeholder proceedings, as well as a detailed assessments of how such meetings are best constructed for maximum gain, see Ryan Hagemann, “New Rules for New Frontiers: Regulating Emerging Technologies in an Era of Soft Law,” *Washburn Law Journal*, Vol. 57, No. 2: 235-263 (Spring 2018), <http://washburnlaw.edu/publications/wlj/issues/57-2.html>.