White Paper: Recommendations on Updating the National Artificial Intelligence Research and Development Strategic Plan

Contributors

PRINCIPAL AUTHORS

Daniel E. Ho, J.D., Ph.D. is the William Benjamin Scott and Luna M. Scott Professor of Law, professor of political science, and senior fellow at the Stanford Institute for Economic Policy Research at Stanford University. He directs the Regulation, Evaluation, and Governance Lab (RegLab) at Stanford, and is a faculty fellow at the Center for Advanced Study in the Behavioral Sciences and associate director of the Stanford Institute for Human-Centered Artificial Intelligence (HAI). He received his J.D. from Yale Law School and Ph.D. from Harvard University and clerked for Judge Stephen F. Williams on the U.S. Court of Appeals for the District of Columbia Circuit.

Jennifer King, Ph.D. is the privacy and data policy fellow at the Stanford Institute for Human-Centered Artificial Intelligence (HAI). She completed her doctorate in information management and systems (information science) at the University of California, Berkeley School of Information. Prior to joining HAI, she was the Director of Consumer Privacy at the Center for Internet and Society at Stanford Law School from 2018 to 2020.

Russell Wald is the director of policy at the Stanford Institute for Human-Centered Artificial Intelligence (HAI), where he is responsible for leading the team that advances Stanford HAI’s engagement with governments and civil society organizations. He is the coauthor of, “Building a National AI Research Resource: A Blueprint for the National Research Cloud,” which is the most comprehensive study on a National Research Cloud to date. Wald has previously held various government relations roles at Stanford University. He is a term member with the Council on Foreign Relations, visiting fellow with the National Security Institute at George Mason University, and a partner with the Truman National Security Project. Wald is a graduate of UCLA.

Daniel Zhang is the policy research manager at the Stanford Institute for Human-Centered Artificial Intelligence (HAI). In this role, he leads the development of the AI Index program and its annual report that aims to measure and evaluate the rapid rate of AI advancement. Daniel also works on the policy team at HAI to develop and manage several policy research programs, building the bridge between technology research and policy communities. Previously, he worked on global AI talent flows and security risks at the Center for Security and Emerging Technology. Daniel holds an M.A. in Security Studies from Georgetown University and a B.A. in Politics & International Affairs from Furman University.

ACKNOWLEDGMENTS

We thank Justin Sherman and Benjamin Bronkema-Bekker for their contribution to this white paper and Jeanina Casusi, Joe Hinman, Nancy King, Shana Lynch, Stacy Peña, and Michi Turner for their help in preparing this publication.
Recommendations on Updating the National Artificial Intelligence Research and Development Strategic Plan

The Stanford Institute for Human-Centered Artificial Intelligence (HAI) offers the following submission for consideration in response to the Request for Information (RFI) by the White House Office of Science and Technology to the Update of the National Artificial Intelligence Research and Development Strategic Plan. Our submission recommends:

- For Strategy 1: Boost non-defense AI R&D budgets, particularly on AI-related infrastructure, to support long-term investments.

- For Strategy 2: Increase support for interdisciplinary and multidisciplinary AI research on human-AI collaboration that expands beyond exclusively technical research.

- For Strategy 3: Please refer to the Stanford HAI letter submitted in January 2022 in response to the White House Office of Science and Technology proposal for an AI Bill of Rights that safeguards the American public against powerful technologies.

- For Strategy 4: Develop appropriate acquisition strategies and update existing procurement regulations to respond to AI procurement and acquisition challenges in the federal government.

- For Strategy 5: Expand government data access to academic researchers to train AI models and develop frameworks for government agencies to evaluate such datasets and their applications in tandem.

- For Strategy 6: Establish a mechanism to evaluate AI models within the exact context of their intended use to ensure safe deployment as well as designate NIST in collaboration with other federal agencies to benchmark AI models in institutional contexts.

- For Strategy 7: Update immigration policies to attract talent in AI and other technical fields as well as develop federal programs to hire AI talent and re-skill civil servants with both technical capacity and institutional knowledge.

- For Strategy 8: Strengthen partnerships with academic institutions and build a framework for a public-university-industry AI R&D ecosystem to drive AI development forward.
Strategy 1: Make long-term investments in AI research.

Recommendation: Boost non-defense AI R&D budgets, particularly on AI-related infrastructure, to support long-term investments.

A long-term commitment to sustained federal research and development (R&D) funding in AI is critical to advance the United States’ leadership in global innovation. The federal government should increase non-defense investment in AI and basic research to strengthen research in critical fields of AI R&D, including healthcare, education, finance, and more, that underpin economic stability and robust growth. Such investment should reflect a multidisciplinary approach, focused on advancing basic and applied R&D, research on AI governance and norm-setting, and supporting research infrastructure with multi-agency collaboration.

Current federal funding for non-defense AI R&D, however, does not meet the needs of the fast-growing AI field. The public non-defense AI R&D budget requested by 25 federal agencies participating in the Networking and Information Technology Research and Development (NITRD) program and the National Artificial Intelligence Initiative in FY 2022 represents an increase of just 8.8 percent over what was spent in FY 2021. In contrast, the National Security Commission on Artificial Intelligence (NSCAI) recommended in its final report to increase public funding for AI R&D at compounding levels, doubling annually to reach $32 billion per year by FY 2026.

Federal long-term non-defense investments—and high return-on-investment basic research funding—can address the challenges the AI innovation ecosystem is currently facing in the United States. For example, the high cost of compute and the lack of access to critical data are hindering efforts by academic researchers to engage in cutting-edge AI R&D. The federal government is lagging behind the private sector on AI development, and federal standards for technical and ethical AI are sorely needed. Long-term public investment in AI-related infrastructure can strengthen AI R&D by supporting a variety of federal initiatives, including the National Artificial Intelligence Research Resource (NAIRR) that aims to expand access to “critical resources and educational tools that will spur AI innovation and economic prosperity nationwide.” Another example of such an initiative is the Multilateral AI Research Institute (MAIRI), recommended by the NSCAI report, that would “facilitate joint efforts to develop technologies that advance responsible, human-centric, and privacy-preserving AI/machine learning (ML) that better societies and allow allies to pool their talents and resources.”

---

7 Schmidt et al., “Final Report.”

Recommendation: Increase support for interdisciplinary and multidisciplinary AI research on human-AI collaboration that expands beyond exclusively technical research.

Intentionally building trustworthy AI that is unbiased and supportive of human flourishing is crucial to ensuring the successful development and deployment of human-centered AI. A key part of that effort requires an interdisciplinary and multidisciplinary approach, involving collaboration from a variety of fields to develop the hardware and software, to understand and design for people’s behaviors and expectations when interacting with AI in different institutional contexts, and to establish policies and regulations to determine human responsibilities as well as the required domain knowledge for various applications. A human-centric approach to AI calls for wide-ranging collaboration among multiple disciplines. Harnessing the potential of AI with rapidly growing capabilities while addressing its impact on existing structural inequalities and biases cannot rely on the voices of computer scientists and engineers alone.

Yet current policies do not necessarily match this need. The National Science Foundation program on algorithmic fairness, for instance, calls for interdisciplinary perspectives while stating that “this program supports the conduct of fundamental computer science research” and requiring the PI to “bring computer science expertise to the research.” But producing fairness-aware algorithms or just understanding the concept of fairness requires knowledge and expertise outside the computer science field to incorporate the social and legal contexts in which AI systems will be deployed. The federal government should expand the support of multidisciplinary AI research for human-AI collaboration to include critical academic fields such as the social sciences, law, ethics, and the humanities—all of which should have a prominent voice in providing the necessary frameworks for understanding AI today and in the future.

Strategy 3: Understand and Address the Ethical, Legal, and Societal Implications of AI.

With respect to recommendations for Strategy 3, please refer to the Stanford HAI letter submitted in January 2022 in response to the White House Office of Science and Technology proposal for an AI Bill of Rights that safeguards the American public against powerful technologies.


Recommendation: Develop appropriate acquisition strategies and update existing procurement regulations to respond to AI procurement and acquisition challenges in the federal government.

Public sector AI can rely heavily on contracting and procurement with external vendors to build up technical capacity. Research shows that almost half of identified use cases of federal agencies’ use of AI came from external sources, with one-third coming from private commercial

---

sources via the procurement process. Compared to the internal sourcing of AI systems that may be more policy-compliant and more accountable, the uses of procured AI in government raise several concerns in terms of AI trustworthiness, transparency, and safety. The federal acquisition regulation (FAR), for example, provides strong IP protection for vendors. But such protections can obscure certain information about the inputs their tools use and how the tools operate behind trade secrecy claims, which in turn prevents appropriate analysis, audit, and testing to ensure the fairness of their use. Moreover, those protections may also create uncertainty for acquisition aimed at the black-box nature of AI systems. FAR makes a clear distinction between rights to “software” and rights to “data,” but AI systems, particularly machine learning (ML), integrate customer software with the new data generated in the process of training, and current procurement policies do not sufficiently address how rights to that data and the resulting AI systems are to be distributed, and under what constraints and conditions.

The federal government should develop appropriate acquisition strategies and update existing procurement regulations to help address some of the public sector’s challenges in evaluating, monitoring, and using AI systems. Specific examples include developing clear standards that call for the disclosure of data and information on the design and operation of contractors’ algorithms, requirements that ensure contractors adhere to ethical AI standards, and testing infrastructures that allow for iterative testing and evaluation.

Strategy 5: Develop shared public datasets and environments for AI training and testing.

Recommendation: Expand government data access to academic researchers to train AI models and develop frameworks for government agencies to evaluate such datasets and their applications in tandem.

While there are publicly available resources for AI development, there is still more work to be done to promote the open and collaborative non-commercial use of training and testing data and environments. For example, access to data resources sufficient for training AI systems is increasingly limited to large private companies, which in turn direct resources toward developing applications with a focus on private profit instead of public interest. Because large platforms have unequaled access to data for AI development, smaller actors, some of which may legitimately lack the financial resources to invest in building training data from scratch, are incentivized to mine the public sphere for data, violating individual privacy expectations and creating privacy risks both for individuals and society at large. At the same time, there are

10 Engstrom et al., “Government by Algorithm.”
significant barriers for interagency and external researchers to access a rich portfolio of public sector data (e.g., employment, healthcare, education). 18

As a starting point, the executive branch can use the NAIRR as an opportunity to make more and better quality government data available to the research community at no cost. 19 In doing so, the federal government should weigh considerations such as privacy, security, and fairness, and it should begin to develop its own frameworks for evaluating such datasets and their applications in tandem, informed by important developments under the Foundations for Evidence-Based Policymaking Act of 2018 and the National Secure Data Service. Considering the lack of a standardized framework to test and evaluate AI models for safety and security, the government should pursue creating a testing environment for this purpose, especially for government-procured AI systems. 20

Strategy 6: Measure and evaluate AI technologies through standards and benchmarks.

Recommendations:

- Establish a mechanism to evaluate AI models within the exact context of their intended use to ensure safe deployment.
- Designate NIST in collaboration with other federal agencies to benchmark AI models in institutional contexts.

Understanding the true in-domain accuracy, or the accuracy of AI systems’ deployment in specific contexts (e.g., in different industries, with different subpopulation groups), is crucial for the federal government to capture the capabilities of the technology and ensure safe deployment. Many current performance evaluations do not comprehensively assess how AI systems would perform in a real-world context. 21 Object recognition systems, for example, are often evaluated against large-scale benchmark datasets to validate their performance, but such datasets are limited in their coverage of non-Western contexts and temporally bounded, capturing only parts of the real world. 22 The same gap is observed in language models where these state-of-the-art systems amplify human bias and discriminate against minority users, and their performance degrades when given out-of-domain text. 23 All told, the accuracy of AI systems in one domain does not automatically translate to its uses in other domains, and changing context can significantly impact performance.

---

22 Raji et al., “AI and the Everything in the Whole Wide World Benchmark.”
The White House should consider a proposal to charge the National Institute of Standards and Technology (NIST), in collaboration with federal agencies that have regulatory oversight for AI-powered products, such as the U.S. Food and Drug Administration (FDA), Consumer Financial Protection Bureau (CFPB), and the National Highway Traffic Safety Administration (NHTSA), to develop improved AI benchmarking protocols. Such benchmarks should explicitly address and incorporate the institutional contexts in which the AI systems are developed (e.g., commercial settings) and deployed (e.g., border control). NIST should also consider how to measure the effectiveness of deploying AI in real-world settings where enabling technologies (e.g., cameras, microphones, computing hardware) and other factors may vary. For instance, one field experiment of an earlier generation of predictive policing algorithms found that models that worked well in the lab did not perform well in the field, failing to reduce crime when used in context. However, it should do so from a human-centered perspective grounded in ethical frameworks, such as privacy by design, and not focus exclusively on identifying technical benchmarks. This requirement suggests a potentially different set of roles and expertise than NIST has had in the past.

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

Recommendations:
- Update immigration policies to attract talent in AI and other technical fields.
- Develop federal programs to hire AI talent and re-skill civil servants with both technical capacity and institutional knowledge.

Future U.S. leadership in AI hinges on the country having the necessary talent-generation process and hiring pipeline—as well as the ability to attract and retain talent that already exists. The country needs individuals who are not only equipped with the skills to build AI systems, but who also know when, where, and how to ask the right questions when such systems pose risks to individuals and society and/or break the law. While not every AI expert has or should have a technical background, understanding the technology is essential to designing and implementing accountable AI initiatives. Such talent could also be a useful tool of accountability, helping design and maintain transparent, auditable, and responsible systems as well as engaging with stakeholders to ensure trustworthiness in AI systems. This could include expediting the hiring process for certain kinds of AI talent, updating immigration policies to attract and retain

---


26 Engstrom et al., “Government by Algorithm.”

technical talent, and expanding opportunities for permanent residency to those with AI and other technical degrees, as well as opportunities for entrepreneurs.

Improving public sector capacity can be essential to accountability and oversight to ensure AI systems are (or are not) built and deployed in ways that promote, rather than degrade, the public interest. Currently, the federal government faces numerous challenges in hiring AI talent, including competing with private-sector salaries and benefits, competing with shorter private-sector hiring timelines, and enticing applicants who face an onerous and often confusing federal hiring process versus an often faster and easier one in industry or civil society. The White House should work on its own, as well as with Congress and with industry partners, to identify the biggest challenges in attracting AI talent to the government and ways to potentially resolve them. Several agencies have initiated such efforts. In 2019, the Office of Personnel Management (OPM) established a classification of information technology (IT) positions to ease the hiring burden of the federal government’s competitive service and the “direct hire” appointing authority for several STEM and IT positions for agencies with critical hiring needs. In 2021, the U.S. General Services Administration (GSA) launched a two-year fellowship aimed at placing early-career software engineers, data scientists, and others with technical skills in federal agencies.

Finally, it bears mentioning that the government should take a long-term view to developing future AI talent and expertise by investing in technical skills development in primary and secondary education through STEM-focused educational initiatives.

**Strategy 8: Expand public-private partnerships to accelerate advances in AI.**

**Recommendation:** Strengthen partnerships with academic institutions and build a framework for a public-university-industry AI R&D ecosystem to drive AI development forward.

The federal workforce to date can still lack appropriate training in AI to use it effectively in public operations and to discharge regulatory responsibilities. A Government Business Council and Accenture survey found that more than 60 percent of federal employee respondents worry about the “lack of technical support and user training” for public AI deployment. An evaluation of the U.S. Customs and Border Protection (CBP)’s facial recognition program used at air exit by the Government Accountability Office (GAO) found that agents on the ground received little

---

training to use a feature of the system. Not only does this undermine agencies trying to carry out their missions, but individuals using AI systems without appropriate training can also create or exacerbate threats to privacy, civil liberties, and even safety. Supporting AI education and research in the university environment can help address some of the government’s talent problems and help fill the talent pipeline for the public sector.

This kind of government-academic collaboration in scientific and technological areas can also fuel innovation. For example, after World War II, the U.S. Department of Veteran Affairs (VA) collaborated with academic institutions (specifically medical centers) to meet the increasing medical needs of returning veterans. The collaboration between academic medicine and the VA helped to revolutionize VA healthcare and spurred innovation in healthcare at many levels, including, for instance, the invention of pacemakers and CAT scan prototypes.

* * * *

As lead authors, we proudly submit this response on behalf of our colleagues and the Stanford Institute for Human-Centered Artificial Intelligence (HAI).

Daniel E. Ho, J.D., Ph.D.
William Benjamin Scott and Luna M. Scott
Professor of Law, Stanford University;
Faculty Associate Director, Stanford Institute
for Human-Centered Artificial Intelligence
(HAI)

Jennifer King, Ph.D.
Privacy and Data Policy Fellow, Stanford
Institute for Human-Centered Artificial
Intelligence (HAI)

Russell C. Wald
Director of Policy, Stanford Institute for
Human-Centered Artificial Intelligence (HAI)

Daniel Zhang
Policy Research Manager, Stanford Institute
for Human-Centered Artificial Intelligence
(HAI)

---