

DAY ONE PROJECT

A National Program for Building
Artificial Intelligence within
Communities

Fernando Delgado

January 2021

The Day One Project offers a platform for ideas that represent a broad range of perspectives across S&T disciplines. The views and opinions expressed in this proposal are those of the author(s) and do not reflect the views and opinions of the Day One Project or its S&T Leadership Council.

Summary

While the United States is a global leader in Artificial Intelligence (AI) research and development (R&D), there has been growing concern that this may not last in the coming decade. China's massive, state-based tech-investment schemes have catapulted the country to the status of a true competitor over the development and export of AI technologies. In response, there have been repeated calls as well as actions by the Federal Government to step up its funding of fundamental and defense AI research. Yet, maintaining our status as a global leader in AI will require not only a focus on fundamental and defense research. As a matter of domestic policy, we must also attend to the growing chasm that increasingly separates advances in state-of-the-art AI techniques from effective and responsible adoption of AI across American society and economy.

To address this chasm, the Biden-Harris administration should establish an applied AI research program within the National Institute of Standards and Technology (NIST) to help community-serving organizations tackle the technological and ethical challenges involved in developing AI systems. This new NIST program would fill a key domestic policy gap in our nation's AI R&D strategy by addressing the growing obstacles and uncertainty confronting AI integration, while broadening the reach of AI as a tool for economic and social betterment nationwide. Program funding would be devoted to research projects co-led by AI researchers and community-based practitioners who would ultimately oversee and operate the AI technology. Research teams would be tasked with co-designing and evaluating an AI system in light of the specific challenges faced by community institutions. Specific areas poised to benefit from this unique multi-stakeholder and cross-sectoral approach development include healthcare, municipal government, and social services.

Challenge and Opportunity

Amidst all the speculative discussion of AI's transformative potential, a chilling effect is curtailing confidence in AI experimentation and development across American society and economy. AI systems are often difficult to understand and at times behave unpredictably, rendering it difficult for human operators to assess their safety or reliability.¹ AI systems have also been shown to perform with less accuracy on minority groups thus exacerbating already existing injustices,² and

¹ Danielle C. Tarraf, William Shelton, Edward Parker, Brien Alkire, Diana Gehlhaus, Justin Grana, Alexis Levedahl, Jasmin Leveille, Jared Mondschein, James Ryseff, Ali Wyne, Dan Elinoff, Edward Geist, Benjamin N. Harris, Eric Hui, Cedric Kenney, Sydne Newberry, Chandler Sachs, Peter Schirmer, Danielle Schlang, Victoria M. Smith, Abbie Tingstad, Padmaja Vedula, and Kristin Warren, *The Department of Defense Posture for Artificial Intelligence: Assessment and Recommendations* (Santa Monica, CA: RAND Corporation, 2019).

² Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," In *Proceedings of ACM Conference on Fairness, Accountability, and Transparency* (2018): 77-91. Clare Garvie, Alvaro Bedoya, and Jonathon Frankle, *The Perpetual Line-Up: Unregulated Police Face Recognition in America* (Washington D.C.: Center on Privacy & Technology, Georgetown Law, 2016).

pose privacy concerns given the type and quantity of data required for effective model development.³

Addressing these challenges is difficult even for the largest and best resourced of organizations. Surmounting them is out of reach for the majority of organizations given the skills gaps and overall costs associated with successful AI development.⁴ The lack of actionable guidance and standards has only added uncertainty to AI adoption, chilling AI innovation across key sectors of American society and economy.⁵ Further, in cases where organizations have pushed forward with AI applications despite the risks, high-profile failures have led to calls for bans on AI technologies from civil society actors.⁶ In aggregate, these various concerns are making it increasingly difficult for many American organizations to prioritize or successfully engage with AI adoption efforts.

There are, however, a variety of niches across American society and economy where AI could play a central role in helping solve deep structural problems for the benefit of the larger public interest. Promising results from pioneering applied research efforts conducted within local organizations demonstrate that AI can be harnessed in hospitals to help reduce clinician errors that lead to misdiagnosis,⁷ in municipal government to protect vulnerable residents from abuse by landlords,⁸ and in human services to enhance child welfare call screening decision-making.⁹ Yet for any area in which there are promising AI applications, there also exist a set of potential pitfalls and growing skepticism regarding their true economic value, trustworthiness, and compatibility with social and economic justice aims.¹⁰

Along those lines, NIST has begun mobilizing around the need for developing AI standards in order to enable the creation of new AI capabilities across American industry and society.¹¹ NIST

³ Cameron F. Kerry, *Protecting Privacy in an AI-driven World* (Washington D.C.: The Brookings Institution, Center for Technology Innovation, 2020).

⁴ Julián Torres Santeli and Sabine Gerson, *5 Challenges for Government Adoption of AI* (Geneva: World Economic Forum, 2019).

⁵ Knut Blind, Soören S. Petersen, Cesare A.F. Riillo, "The Impact of Standards and Regulation on Innovation in Uncertain Markets," *Research Policy* 46, (2017): 249-264; Will Hunt, *The Flight to Safety-Critical AI: Lessons in AI Safety from the Aviation Industry* (Berkeley: Center for Long-Term Cybersecurity, 2020).

⁶ Daniel E. Ho, Emily Black, Maneesha Agrawala, and Fei-Fei Li, *Evaluating Facial Recognition Technology: A Protocol for Performance Assessment in New Domains* (Palo Alto: Stanford Institute Human-Centered Artificial Intelligence, 2020).

⁷ Nan Wu, Jason Phang, Jungkyu Park, Yiqiu Shen, Zhe Huang, Masha Zorin, Stanislaw Jastrzębski et al. "Deep neural networks improve radiologists' performance in breast cancer screening," *IEEE Transactions on Medical Imaging* 39, no. 4 (2019): 1184-1194.

⁸ Teng Ye, Rebecca Johnson, Samantha Fu, Jerica Copeny, Bridgit Donnelly, Alex Freeman, Mirian Lima, Joe Walsh, and Rayid Ghani, "Using Machine Learning to Help Vulnerable Tenants in New York City," In *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies* (2019): 248-258.

⁹ Allegheny County. Allegheny Family Screening Tool. Accessed December 14, 2020, <https://www.allegHENYcounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx>.

¹⁰ Mariano Florentino-Cuéllar, David Freeman Engstrom, Daniel E. Ho and Catherine Sharkey, *AI's Promise and Peril for the U.S. Government* (Palo Alto: Stanford Institute for Human-Centered Artificial Intelligence, 2020); Madeleine Elish and Elizabeth Watkins, *Repairing Innovation: A Study of Integrating AI in Clinical Care* (New York: Data & Society, 2020).

¹¹ National Institute of Standards and Technology, *U.S. Leadership in AI: A plan for Federal Engagement in Developing Technical Standards and Related Tools*, August 2019, https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf.

has drafted reports on AI terminology¹² as well AI interpretability and explainability¹³ that have been made available for public comment and are set to be finalized in the new year. In 2020, NIST also began holding workshops on key AI issues including bias, trustworthiness, and explainability.¹⁴ Prior to these most recent efforts, and for well over a decade, NIST has played a central role in establishing standards and enabling industry innovation for early waves of AI-related technologies such as automated facial recognition¹⁵ and natural language processing.¹⁶

What is needed now is a targeted hyperlocal program to stimulate innovative applied research projects across a broader swath of American society and economy. The most effective strategy would focus on organizations and communities that are in need of novel solutions yet have limited AI access due to lack of resources, coordination, and vision. This program would seek to incentivize industry and academic researchers to responsibly apply state-of-the-art AI techniques to complex real-world situations. This program would also serve as a catalyst to educate and empower practitioners on the ground to take ownership of AI tools and process, equipping them with the experience to carry forward the maintenance and evolution of applied AI practice in their domains.

There are pioneering examples to draw inspiration from. Across law and medicine sectors, an interdisciplinary and practitioner-driven approach to AI innovation has led to the development of robust systems capable of addressing structural problems facing mission-critical clinical decision-making.¹⁷ These efforts represent a distinctively American approach to innovation adept at working across traditional boundaries and generating new forms of technical know-how and economic value within communities. These efforts emphasized collaborative experimentation, learning, and deliberation as part of the innovation process. And they were anchored to the ultimate goal of developing working systems to meet the specific needs of the targeted organizations and the populations they serve. We must leverage these unique real-world precedents to devise a federal program that incentivizes and bolsters the building of AI systems and capabilities within American communities.

¹² Elham Tabassi, Kevin J. Burns, Michael Hadjimichael, Andres D. Molina-Markham, and Julian T. Sexton, A Taxonomy and Terminology of Adversarial Machine Learning (DRAFT NISTIR 8269), October 2019, <https://doi.org/10.6028/NIST.IR.8269-draft>.

¹³ P. Jonathan Phillips, Carina A. Hahn, Peter C. Fontana, David A. Broniatowski, and Mark A. Przybocki, Four Principles of Explainable Artificial Intelligence (DRAFT NISTIR 8312), August 2020, <https://doi.org/10.6028/NIST.IR.8312-draft>.

¹⁴ National Institute of Standards and Technology. NIST AI Workshop Series. Accessed December 14, 2020, <https://www.nist.gov/artificial-intelligence/nist-ai-workshop-series>.

¹⁵ Facial Recognition Technology, Before the Committee on Homeland Security, 116th Congress (2020) (statement of Dr. Charles H. Romine, Director, Information Technology Laboratory, National Institute of Standards and Technology, United States Department of Commerce).

¹⁶ National Institute of Standards and Technology. Text REtrieval Conference (TREC). Accessed December 14, 2020, <https://trec.nist.gov>.

¹⁷ Fernando A. Delgado, "Sociotechnical Design in Legal Algorithmic Decision-Making," In Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing (2020): 111-115; Mark Sendak, Madeleine Elish, Michael Gao, Joseph Futoma, William Ratliff, Marshall Nichols, Armando Bedoya, Suresh Balu, and Cara O'Brien, "The Human Body is a Black Box: Supporting Clinical Decision- Making with Deep Learning," In Proceedings of ACM Conference on Fairness, Accountability, and Transparency (2020): 27-30..

This program calls upon the Federal Government's unique power to convene across America's key innovation players: academia, industry, the public sector, and civil society. This program would bolster technological innovation and capability development outside of the traditional research lab and directly into practice across American organizations, all while developing the experience, knowledge base, and results required for targeted guidance and standards development. It would constitute a uniquely pragmatic and decentralized American approach to AI capability development, rooted in high priority domestic concerns, challenges, and goals.

Plan of Action

The Biden-Harris administration should create a new program within NIST dedicated to funding and coordinating applied AI research initiatives rooted in local community organizations. The proposed new program would fund and coordinate projects within specific organizations who are working with AI researchers to assess the viability of integrating AI techniques into their local decision-making workflow. The first goal the program would be to translate fundamental AI research into the development of innovative AI-based systems. The second goal would be to ensure that these systems are aligned to local needs and that they maximize social and economic benefits while minimizing risks.

This program would fund several research teams each composed and co-led by technology researchers and community practitioners (i.e., clinicians, administrators, advocates). This program would fund three site-specific projects each year, each of which would be tasked with devising an AI system to address the specific challenges faced by the organization and the community it serves. The program would provide funding for the time worked by principal investigators and their assistants as well as the data infrastructure and program management resources necessary for team coordination and information-sharing. Funding would operate via direct grants to participants and assumed costs related to the salaries of dedicated agency staff, shared technical infrastructure, and ongoing workshop events. Each project would be budgeted at \$2 million annually to allow for well-staffed teams composed of AI researchers and domain practitioners to participate in a dedicated and focused manner. With three concurrent projects running across four years, and an incremental agency overhead expenditure of \$1 million annually to oversee the program, the total funding cost would be \$28 million.

NIST is especially well-suited to house this new program office given its already central role in federal AI R&D strategy and its established links to other AI R&D key players like the National Science and Technology Council (NSTC), the Office of Science and Technology Policy (OSTP), and the AI Research and Development Interagency Working Group. Additionally, NIST's own foundational and principles-oriented AI research initiatives would naturally complement the application-focus of this program. Most importantly, NIST already has an exceptional track record

DAY ONE PROJECT

in bridging the chasm between theoretical research and useful applications across various areas of American technology including internet search, genomics, and legal informatics.¹⁸

¹⁸Brent R. Rowe, Dallas W. Wod, Albert N. Link, and Diglio A. Simoni, Economic Impact Assessment of NIST'S Text REtrieval Conference (TREC) Program, July 2010, <https://trec.nist.gov/pubs/2010.economic.impact.pdf>.

Frequently Asked Questions

What is the difference between fundamental AI research and applied research?

While fundamental research in AI seeks to advance the state of knowledge and theory on algorithmic techniques generally, applied AI research focuses on the design, fit, and maintenance of systems leveraging algorithms for particular users and communities. As such, applied AI research is necessarily interdisciplinary requiring not only computational expertise but also design, social scientific, and domain knowledge in its execution.

Who exactly are the full set of stakeholders envisioned in this proposal?

Taking the example of clinical decision-making in a hospital setting, in addition to academic and industry computer science and systems design researchers, a project team would also collaborate with medical doctors, nurses, and hospital administrators associated with an organization that is looking to integrate AI into its workflow. The research would be co-led by researchers and practitioners while also looping in additional representatives from government, medical professional guidance bodies, and civil society groups who specialize in tracking the state-of-practice in the chosen area.

How will you incentivize participation from leading AI researchers?

The holistic and applied framing of this program would attract leading researchers interested in tackling complex human factors and sociotechnical problems facing AI research. Additionally, the program should recruit and incentivize participation from teams already working on related research funded by the National Science Foundation (NSF), in particular research being undertaken through its National AI Research Institutes program. Since reward structures across academic and industry research teams favor publications, the program should facilitate publication of research team's work in leading scientific venues. Also, housing this program within NIST, a well-known and respected standard setting body, researchers will be motivated to contribute to the effort since their participation could result in them helping shape standards for their area of practice.

How is this different from Kaggle or other common task challenges?

Kaggle is a platform that allows data scientists and machine learning engineers to find and publish data sets, explore and build models, and collaborate with each other. There are indeed some similarities with this program and other AI task challenges venues (i.e., multiple research teams, shared data sets, creation of benchmarks). Yet, the focus of this program is on fostering a collaborative and exploratory research exercise, where in addition to examining the performance of the underlying technology, attention is centrally paid to how people on the ground interact with the technological components of the system. This sociotechnical orientation

DAY ONE PROJECT

is necessary to achieve suitable performance for real-world tasks and also for mitigating potential negative downstream impacts. In other words, this framework builds upon the successes of common task challenges, but radically transforms the questions and goals of these technical community research efforts.

DAY ONE PROJECT



About the Authors

Fernando Delgado is a PhD candidate in Information Science at Cornell University. Prior to commencing his doctoral studies, Fernando worked at H5, a pioneering firm in the field of legal technology designing and deploying algorithmic systems for fact-finding in civil litigation.

His current academic research focuses on elaborating and refining design, evaluation, and governance frameworks for automated decision systems in high-stakes domains. He draws on theory and methods from the social science of technology and law, as well as the computational fields of information retrieval and machine learning. His research is supported by the McNair Scholars Program, the MacArthur Foundation program on Technology in the Public Interest, and the Russell Sage Foundation initiative on Computational Social Science.



About the Day One Project

The Day One Project is dedicated to democratizing the policymaking process by working with new and expert voices across the science and technology community, helping to develop actionable policies that can improve the lives of all Americans, and readying them for Day One of the next presidential term. For more about the Day One Project, visit dayoneproject.org.