PUBLIC SUBMISSION

Received: May 29, 2025
Tracking No. mb9-vn8q-tq97
Comments Due: May 28, 2025 **Submission Type:** Web

Docket: NSF-2025-OGC-0001
NITRD_FRDOC_0001

Comment On: NSF-2025-OGC-0001-0001
Request for Information: Development of a 2025 National Artificial Intelligence Research and Development Strategic Plan

Document: NSF-2025-OGC-0001-DRAFT-0264
Comment on FR Doc # 2025-07332

Submitter Information

Organization: Indiana University

General Comment

This is an update to a previous comment submitted by Indiana University. The previous comment did not have page numbers. The reference number of the previous submisison is mb9-v66r-aqff. Please replace the previously submitted .docx file with the attached .pdf file.

Attachments

IU_AI_RFI_Response_FINAL

Indiana University's response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development (R&D) Strategic Plan

As a leader in the national AI research community, Indiana University (IU) has multiple perspectives on the recent request for information from the Networking and Information Technology Research and Development National Coordination Office. IU faculty are at the forefront of AI research, from fundamental AI research in the IU Luddy School of Informatics, Computing, and Engineering, to the application of AI technologies in business, medicine, and the physical sciences. IU is also one of the inaugural members of, and a resource provider for the National AI Research Resource (NAIRR) Pilot[1], is participating in two NSF AI institutes[2,3] and leading a planning grant for another.[4] IU Is also a founding member of NIST's AI Safety Institute Consortium, and was tasked by NSF to map AI curricula across the nation as it pertains to cybersecurity.[5] From these perspectives, IU brings a unique voice to this request for input on updates to the 2023 update of the National Artificial Intelligence Research and Development Strategic Plan (NAIRDSP).

The initial 2016 NAIRDSP contained seven strategic objectives, to which the 2019 update added an eighth (Strategy 8: Expand Public–Private Partnerships to Accelerate Advances in AI) and the 2023 update added a ninth (Strategy 9: Establish a principled and coordinated approach to international collaboration in AI research). While progress has been made in many of the strategic objectives, there remains much work to do. Additionally, recent developments in large language models (LLMs) and other foundational models have expanded the need for the objectives surrounding ethics, safety, and security called for in strategies 3 and 4 of the 2023 update. Our response is organized around the strategic goals from the 2023 update, specifically *Strategy 5: Develop Shared Public Datasets and Environments for AI Training and Testing*, and *Strategy 6: Measure and Evaluate AI Systems through Standards and Benchmarks*. It is IU's view that these two strategies are the linchpins in achieving success in the other strategies set forth in the 2023 update. While all of the strategies outlined in the 2023 update are relevant in today's AI landscape, the fact that they are presented as nine relatively independent areas with some modest attempts at interlinking them is the biggest shortcoming of the overall strategic plan. We propose identifying and grouping the strategies according to their primary objective and relation to the other strategies.

[1] https://nairrpilot.org/
[2] https://icicle.osu.edu/
[3] https://engageai.org
[4] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2020194
[5] https://www.nist.gov/aisi/artificial-intelligence-safety-institute-consortium/aisic-members

Strategies 1-4 are focused on the **outcomes, features, and capabilities** of AI systems created from R&D efforts. With strategy one being recast as expanded capability or functionality, strategy two focused on human/AI interaction, strategy three focused on legal and ethical AI, and strategy four focused on safe and trustworthy AI. All four of these capabilities are important aspects of AI research, and special attention should be paid to the areas that are lagging behind the others. In particular, expenditures to research safeguards to produce safe, ethical, trustworthy, and explainable AI should be made to allow these strategic foci to gain parity with the others. In the absence of regulation, few industry R&D efforts will be targeted at these areas.

Strategies 5 and 6 focus on the **infrastructure and foundational elements** from which innovations from the R&D efforts in strategies will come. Even the most theoretical research topics in AI (e.g. ethics, the theoretical limits of AI, etc.) require access to data, computational platforms, software, and professional staff to support research programs.

Strategies 7-9 focus on the **integration of AI research** with a variety of different sectors, including integration with the current and future workforce and understanding how to cultivate an AI literate workforce; integration with industry and other non-federal entities; and integration with international partners who are focused on similar principles and strategies around advancing AI.

Viewing the nine strategies outlined in the 2023 update in this structure, instead of nine equal initiatives, yields a structure with strategies 5 & 6 acting as a foundation for strategies 1-4 and strategies 7-9 woven throughout the various goals. This perspective underscores the impact of strategies 5 & 6 in achieving the goals of the other seven strategies, all of which are paramount to the United States maintaining and extending its global leadership in the AI space.

Strategy 5: Develop Shared Public Datasets and Environments for AI Training and Testing

The foundation for all AI research and development activities is a strong cyberinfrastructure ecosystem, and IU has been working toward strengthening the AI and general scientific cyberinfrastructure ecosystem for more than two decades. We use here the definition of cyberinfrastructure proposed by Stewart et al (2010), namely, “Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high-performance networks to improve research productivity and enable breakthroughs not otherwise possible.” IU has acted as a resource provider in the national cyberinfrastructure ecosystem beginning with the Teragrid[6] and

[6] https://hdl.handle.net/2142/43874

accelerating with the deployment of Jetstream[7] in XSEDE[8], the NSF's first production cloud system for scientific research. Since then, IU has continued to expand on that success, providing Jetstream2[9], a GPU powered, AI focused cloud computing resource for both ACCESS[10] and the NAIRR Pilot. In addition, IU is a partner in the ICICLE AI collaboration, which is focused on creating the next generation of cyberinfrastructure for AI and democratizing access to all AI researchers.

While the first year of the NAIRR Pilot has been very successful, given the limited funding allocated to it, there are a number of lessons that have been learned and can be applied in the implementation of a full NAIRR realization. Resources for NAIRR have either been through public or private contributions or via agencies (e.g. NSF, DOE, and NIH) re-allocating existing funds to extend or expand existing services. Several of these items have already been alluded to in the 2023 update of the strategic plan, while others are wholly new. Here we highlight and reiterate those found in the 2023 update and expound on the new findings.

- Access to data: Nearly all AI research is reliant on large amounts of data, i.e. big data. The 2023 update cites the large amounts of data available in federal data sets and highlights the challenges of applying FAIR (Findable, Accessible, Interoperable, and Reusable) to those data sets. Additionally, locating the data near the computational resources that are made available in a platform like NAIRR can be a challenge. These challenges have been observed in the NAIRR Pilot and could be addressed by provisioning replicas of data sets that are of particular interest for AI researchers at NAIRR computational facilities. This would make these data more readily available to researchers. However, the challenge remains that many of these data sets were not collected with AI research in mind, and so may be missing essential metadata or may be lacking in some aspect for AI research. In addition to easing the access for already existing data sets, there should be a focus on obtaining additional data sets that are curated *specifically for* AI research.
- Access to large scale *and modest scale* computing resources: The NAIRR Pilot has had great success in providing GPU and CPU resources to researchers from institutions throughout research and higher education. While the 2023 update calls for shared large-scale hardware, the wide variety of research areas pursuing AI R&D and education has shown that a spectrum of resources are necessary, from smaller scale single GPU

[7] https://doi.org/10.1145/2949550.2949639
[8] https://hdl.handle.net/2142/96220
[9] https://doi.org/10.1145/3437359.3465565
[10] https://access-ci.org/

resources for education and lightweight inference tasks to massive numbers of GPUs for training of foundational models. In fact, the NAIRR Pilot has yet to deploy resources at a scale that are required for transformational research such as the production of large-scale foundational models across many science disciplines. These capability scale resources are needed to make progress in a number of areas, but the progress that has been made using the more modest resources of the NAIRR Pilot should be continued as well. In addition to a broad spectrum of scales for more traditional resources (CPU and GPU resources), access to AI-specific hardware, such as the devices from Cerberas and Samba Nova available through the NAIRR pilot, should be increased. Most of the studies using new AI-specific hardware or accelerators have shown positive results, (c.f. Santos et al. 2024)[11] but have been exploratory. Only by making these resources more widely available will we be able to see their adoption and assess their utility in a variety of science disciplines.

- Access to models and their provenance: Both data and computation are essential elements of most, if not all, AI research. With the advent of LLMs like ChatGPT [ref], research teams are pursuing a number of avenues that leverage open weight models (e.g. Meta's Llama [ref]). An open weight model differs from open source in that the creator of the model makes openly available a weight file, but not details about how the model was created or the provenance of the data used to create the model. This is an extension of the already existing and widespread problem in AI research of not providing model metadata and provenance. The Hugging Face repository[12], the primary model repository for AI research, is full of models with little to no provenance information. As generative AI becomes more widely used in a variety of research settings, the fact that most, if not all, credible LLMs are either closed or open weight creates several challenges for reproducibility, intercomparison of model performance, and model extension. Efforts should be made to create and provide foundational models that are fully open, including the data that have been used to train the models, the algorithms used for training, and the training methodology. Providing a framework to catalog and make these artifacts available is of utmost importance and is another key piece of AI cyberinfrastructure.
- Support and facilitation from cyberinfrastructure professionals: A key component in the above definition of cyberinfrastructure is the professional staff needed to provide resources for research. These staff include system

[11] https://doi.org/10.1109/SC41406.2024.00014
[12] https://huggingface.co

administrators, cybersecurity experts, and user support staff. But perhaps one of the more overlooked roles in the AI landscape is the role of research facilitator. A research facilitator is someone who helps to translate the ideas and direction of the AI or domain science researcher into an implementation for a given set of hardware and software tools that are available for that research endeavor. This work is often a challenge that lies outside the area of expertise of even senior AI researchers. The software frameworks used to train and use AI models are often quite complex, and there are many differences between hardware implementations. For domain scientists or other subject matter experts, navigating the various possibilities of AI algorithms and hardware infrastructures can be impossible without help from a facilitator. A recent example comes from a science education researcher in the IU School of Education who was pursuing the idea of training an LLM to mimic the behavior of students to assist in teacher training. Without support from research facilitators with expertise in model fine-tuning frameworks and how to deploy those workflows to a supercomputer, the research team could not make progress.

- Interoperation between federal agencies and industry partners: Another observation from the NAIRR Pilot is while the NSF has done an outstanding job of facilitating access to computational and storage resources across a very broad portfolio of government agencies and public/private partnerships with no additional agency funding, there have been a number of challenges in presenting the NAIRR as a cohesive resource for AI research. This is partly due to the inherent differences in how computational resources are operated and provided across diverse participants like the Department of Energy, NSF resource providers, the National Institutes of Health, and various commercial entities, including NVIDIA, OpenAI, Microsoft, Amazon, and Google. It is also due to the lack of targeted funding and direction to provide a more cohesive front door to these resources. In order to present a unified AI mission across the nation, it is imperative to have a framework that coordinates access to resources for R&D across government, industry, and academic agencies. These agencies sharpen our security, competitive edge, and innovation now and for future generations across the nation. Initiatives like the NAIRR Pilot are central to the creation of the processes and structures necessary to facilitate access to resources from industry partners like NVIDIA, national labs (ANL), and research institutes that, in turn, drive these missions forward. Failure to fund these initiatives may hamper the advancements that have been enabled by NAIRR.

Each of these key cyberinfrastructure elements: data, computational power, models and algorithms, cyberinfrastructure professionals, and a coordinating framework, is needed to enable AI research and to serve as a foundational platform for the R&D that the US needs to achieve its AI potential. This includes funding the NSF at the FY2023 funding level requested by the CNSF[13] ($9.9B), funding the full NAIRR project as recommended by the NAIRR Taskforce[14] ($2.6B), and, as Sen. Todd Young (R-IN) recently stated at an Indiana Business Journal event, "[Congress] should fund the science portion of the chips and sciences act." These initiatives are fundamental structures that continue to provide research strategies that have allowed the United States to be at the forefront of science and technology while also training a workforce that is necessary for national security and global competitiveness.

Strategy 6: Measure and Evaluate AI Systems through Standards and Benchmarks

As it is with any practice, setting standards, benchmarks, and best practices establishes a baseline in ensuring quality, reliability, and interoperability across the infrastructure required for AI systems. When an accepted standard framework for measuring and evaluating these systems exists and has been adopted by the global community, this in turn fosters competitive innovation by instituting benchmarks that industries and organizations can measure their systems against. The 2023 update outlines the shared concerns of the lack of standardization of benchmarking across AI systems, which undermines our ability to measure progress, protect our systems and plan towards future AI capacity and capabilities. The shortcomings in standardized benchmarks span hardware and software frameworks, and apply to performance, efficiency, accuracy, and safety.

- Establishing AI Technology Benchmarks
  The lack of a widely accepted suite of benchmarks that define the capabilities of a system and set the requirements for a system to achieve its stated goals and mission leads to an environment without a means to evaluate and assess AI systems for safety, efficiency and performance. Establishing these benchmarks solidifies credibility of published performance results in the world market and facilitates R&D efforts to attain the results they are striving for.
  Moreover, this response to the RFI would like to emphasize the opportunity for the United States to play a leading role in determining and setting these standards and best practices in the global market. Through the experience gained from rigorous acceptance reviews for numerous cluster and cloud deployments, IU is positioned to contribute to developing and sharing standards

[13] https://www.cnsf.us/letters-statements
[14] https://www.ai.gov/wp-content/uploads/2023/01/NAIRR-TF-Final-Report-2023.pdf

for benchmarking systems for acceptance. Further support is necessary to continue, improve, and expand these efforts will propel competitive innovation in AI R&D in both public and private sectors.

- Increasing the Availability of AI Testbeds
  With the rapidly evolving landscape of AI systems, the deployment of and access to AI testbeds are necessary components towards the adoption and advancement of a broad spectrum of benchmarks that set the requirements and guidelines for acceptance of the variety of large-scale infrastructures. By providing access to standardized, secure operational environments, we ensure the interoperability across AI resources and create a path towards scaling up R&D activities to larger AI systems that adhere to the same standards. Funding to enable the installation, support and maintenance of these testbeds will safeguard our nation's position at the front of the advancements in AI in addition to giving us the ability to characterize the capabilities of available AI systems and leverage that knowledge to drive forward future infrastructures.
  Data centers across the nation, including the one located at IU's flagship Bloomington campus, can provide insight into what our current capacity is for housing or expanding to house AI testbeds at sites close to researchers and scientists and that have created processes and frameworks for onboarding users onto experimental systems. Funding the deployment of these testbeds at these centers and the frameworks that provide access to them like NAIRR are critical requirements for ensuring a standard for powering R&D AI activities in our country.
- Engaging the AI Community in Standards and Benchmarks
  Without the adoption of standards across communities, several risks are introduced, not just to our national security but also to our nation's position as a leader in AI innovation and our ability to influence the trajectory of the characteristics and capabilities of AI systems that we see increasingly applied in all facets of industry and everyday life. Adoption will need a framework for driving compliance to the standards in communities. Dedicated support is necessary for engaging these communities to create a unified voice and detailed plan for the promotion and adoption of these benchmarks and best practices. As an R1 university that has an established network of collaborators in research, academia, government and industry, IU and similar institutions across the nation should be empowered and supported to drive this mission forward by leveraging these relationships and our credibility as a research institution. For example, the Nobel Prize-winning Ostrom Workshop has done pioneering work on AI governance and is working with other leading institutions, including UC Berkeley,

on frameworks and standards to guide the evolution and adoption of new AI technologies.[15]

Importance of capabilities strategies

While the foundational nature of strategies 5 & 6 makes them essential to AI R&D, they are just the first step. The capabilities strategies (1-4) are equally important in fueling discovery and ensuring that AI capabilities are developed and deployed in an ethical, safe, and trustworthy manner. Funding is needed for investigations in improving current AI hardware and creating specialized AI hardware, more efficient algorithms, and software frameworks, and applying AI techniques across a range of scientific disciplines. Studies on how to identify, prevent, and/or manage biases in AI systems are critical as well. Investigations into ways to create AI systems that are both transparent and explainable have become even more important as AI models have become ever more complex and are beginning to be embedded into complex workflows. Research on how multiple AI systems interact with each other in agentic workflows is an emerging need to maintain safe and trustworthy systems.

Equally vital is the infrastructure required to support the rapid growth of these AI capabilities. As AI research expands, its influence will extend across government and industry, making secure, accredited facilities essential for handling everything from sensitive data to classified Department of Defense data. Recognizing this need, Indiana University has invested in its Applied Research Center, designed to safeguard the government's most protected information. These facilities not only enhance AI research and development but also provide strategic interagency benefits, maximizing funding efficiency.

From AI institutes[16,17], to machine learning driven studies in confidential computing[18], to structural protein analysis with tools like Alphafold[19], to improving teacher efficacy[20] and student AI literacy[21], NSF awards have pushed the forefront of AI R&D. It is essential that the NSF continue to fund investigations at all levels from novel innovations at small team level, to collaborative efforts like the AI institutes, to a broad national cyberinfrastructure ecosystem like NAIRR. IU will continue to advance AI research at all

[15] https://ostromworkshop.indiana.edu/research/data-management/index.html
[16] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2020194
[17] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2112606
[18] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2207231
[19] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2441717
[20] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2334631
[21] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2417275

these levels and strongly urges the NSF and other federal agencies to continue funding research in these areas.