PUBLIC SUBMISSION

Received: May 29, 2025
Tracking No. mb9-u8e2-b39k **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-0258
Comment on FR Doc # 2025-07332

Submitter Information

Government Agency Type: State
Government Agency: University of Arizona

General Comment

See attached file(s)

Attachments

University of Arizona Response to RFI on the Development of a 2025 National Artificial Intelligence Research and Development Strategic Plan

University of Arizona response to the Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development (R&D) Strategic Plan, Docket ID No. NSF-2025-OGC-0001

This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in developing the 2025 National AI R&D Strategic Plan and associated documents without attribution.

The United States is an unrivalled global leader in AI due to a strong innovation ecosystem involving industry, academia, and government. Continued government R&D is essential to maintain this leadership. The following key areas articulate research needs and development challenges that will benefit from federal government prioritization, as they are less likely to receive sufficient private sector investment but are critical for national interests and for maintaining U.S. leadership in AI.

1. Advancing Foundational AI Research: Understanding Learning Systems

The 2025 National AI R&D Strategic Plan should prioritize support for foundational research in machine learning, with emphasis on the inner workings of neural networks. Despite rapid progress in AI capabilities, a fundamental lack of scientific understanding and clarity remain around how these systems operate at scale.

A. Understanding and Improving Learning Systems

We recommend strategic investments in the following efforts essential for building more transparent, robust, secure, and trustable AI systems:

- Define and trace the causes of hallucinations, reverberations, and related failure modes in large models.
- Develop tools to explain and interpret these behaviors and their underlying mechanisms.
- Extend explanation and interpretability capabilities to encompass the learning processes themselves.
- Design interactive visual analysis tools that support both failure diagnosis and exploration of learning dynamics.
- Identify architectural or training innovations that mitigate hallucinations, reverberations, and related failure modes at the system level and promote more interpretable and transparent learning processes.
- Ultimately, improve AI reasoning and interpretability through analysis, visualization, and interventions.
- Develop diagnostic and prognostic tools to identify and predict system dysfunction on-line and a priori.

A critical challenge is the need to reduce the time and energy required for training machine learning models. With deep neural networks now characterized by enormous numbers of nodes and weights, training processes have become exceedingly resource intensive. Concurrently, another key priority involves reducing the volume of data necessary for effective model training. The prevailing big-data-centric approach, which indiscriminately uses massive datasets, proves highly inefficient, as the resulting trained models often yield only moderate performance improvements despite consuming extensive computational resources. The focus moving forward should be on data cleaning and preconditioning, based on data science principles, and developing new data science principles for this.

B. Architectural and Theoretical Innovations

Basic research is needed into alternative deep learning models and entirely different architectures. This includes, but is not limited to, neuromorphic computing, biocomputing, neuro-biomimicry, provable and explainable reasoning, causal models, AI integration with real-world simulation, inductive logic, genetic algorithms, Bayesian models, probabilistic programming, human-guided AI collaboration, ethics, policy development, and many other fields.

Developing a versatile quantum learning machine hardware platform is of paramount importance. Such a platform must achieve substantial reductions in energy consumption while delivering exceptionally high information processing bandwidth. This approach leverages quantum computing's inherent capabilities, enabling highly efficient, low-energy processing beyond classical limitations.

Equally important is providing mathematically rigorous proofs addressing two fundamental areas: first, the universal approximation capability of various machine learning model classes, which confirms their theoretical potential; and second, a deeper understanding of computational “expressivity” inherent to specific hardware or software algorithms. These theoretical advances will guide practical innovation, ensuring optimal alignment between computational architectures and algorithmic strategies.

2. Promoting High-Impact, Domain-Specific Applications of AI

The national strategy should include targeted investments in research that adapt AI to critical domains, where its use has the potential to transform outcomes. For AI to be truly effective and trusted in these critical domains, it must be firmly grounded in scientific principles. Priority areas include:

- **Medical and scientific instrumentation**: Enhancing diagnostics, monitoring, and operational automation.
- **Healthcare:** Empowering provider-patient interactions for improved outcomes, prevention, enhanced patient satisfaction/experience, and cost savings.
- **Business, financial, and clinical settings**: Producing reliable summaries and insights from complex data streams.
- **AI for science:** Unlocking scientific discovery through causal inference and model-driven reasoning.
    - **Quantum technologies:** Enhancing AI to model, control, and accelerate advances in quantum systems and computing.
    - **Materials science:** Accelerating discovery and characterization of novel compounds and processes.
    - **Fusion energy:** Applying AI to accelerate progress in complex, data-intensive energy systems.
    - **Critical minerals and mining**: Using AI to improve exploration, extraction efficiency, and environmental stewardship.

  - **Astronomy**: Integrating machine learning, large-scale data mining, and advanced computer vision to process and interpret massive, multi-modal astronomical datasets.
- **Space and national security:** Advancing trustworthy, explainable AI for autonomous systems and strategic situational awareness.
- **Intelligence engineering:** Integrating AI methods into (existing) systems.
- **AI-human collaboration**: Designing AI to augment human expertise, including developing human-guided AI.

Human-AI co-creation and collaboration should be supported through novel computational architectures that enable AI systems to reason about complementary human-machine strengths, incorporate multi-modal communication paradigms (spanning language and visual interfaces), and integrate with automation for end-to-end productivity solutions.

Strategic investment should also support the development of evaluation methodologies that assess innovation, learning, and knowledge transfer in human-AI systems.

3. Enhancing Trustworthiness Without Hindering Innovation

Improving the trustworthiness, reliability, and explainability of AI systems must remain a key strategic objective. However, overly restrictive approaches may stifle innovation and delay beneficial deployment and adoption.

We recommend:

- Developing tools and methods to monitor information veracity and identify misinformation presence, growth, and expansion, with alerts for system mechanisms for decline or decay.
- Developing methods and tools for assuring reproducibility, with a reproducibility index.
- Developing standards and metrics to evaluate overall trustworthiness of AI platforms in an automated manner, with the ability to identify changes and alert users to deviations and allow user communities to define acceptable use criteria and validation standards. Ultimately, the most important metric is not just prediction quality in isolation, but the ability of AI models to support trustworthy action -- especially in data-scarce, unstable, or safety-critical regimes.

4. Data Management Strategies to Enhance Innovation

The utility of data increases when it is used to create models and when disparate data sets are combined, enabling new discoveries. While open access to data will enhance innovation, this policy can be exploited by bad actors who profit from the availability of data but do not contribute to the data ecosystem.

Additionally, we have a number of regulatory frameworks aimed at controlling access to data. Some of these frameworks, like HIPAA and GDPR, are intended to protect an individual's right to privacy. Others, like ITAR, are intended to protect defense-related information that could be used by an adversary. While well intended, polices like the GDPR can have unintended consequences; it has been shown that the GDPR was a much greater relative burden for small companies than for larger companies. This can have a negative impact on innovation.

The national AI strategy should include support for research that explores controlled access to data that is informed by the inherent value of the data, the rights of individuals, and potential impact on innovation.

5. Strategic Research Gaps Unaddressed by Industry

- Industry's focus on transformer architectures has driven many breakthroughs, but may hinder exploration of potentially superior, undiscovered architectures. Government investment should support non-transformer model research.
- Current industry efforts emphasize jack-of-all-trades models (e.g., large language models) that are acceptable for chatbot-like tasks but are poorly suited for high-risk environments, such as healthcare. Government research should prioritize developing domain-specific algorithms tailored to critical applications.
- Industry is often driven by easily available data sources (e.g., web pages, captioned images), limiting their models' utility in domains lacking large datasets. Federal support is needed for curating and digitizing domain-specific data sets, making them AI ready for developing robust AI systems capable of functioning with small or limited data.

6. Educational Aspects of AI Adoption

A continuum of skill levels is needed for people to efficiently, safely, and reliably use AI. Engagement should advance from basic AI literacy to AI fluency, empowering individuals not just to understand AI but to critically engage with its development, applications, and societal impacts. Strategic priorities should reflect this diversity by supporting the following areas.

A. Broad AI Literacy and Public Engagement

- Training for K-12 teachers, librarians, and other knowledge workers to help the general population use AI effectively and responsibly.
- Research on new approaches to increasing AI literacy and achieving AI fluency by expanding training in ethics and critical use of AI technologies.
- Tools integrated into workplace systems to support learning-by-doing in authentic contexts, enabling users to develop AI fluency while engaging in meaningful tasks.
- Access to baseline AI infrastructure to teach, train, explore, and use AI-based tools (e.g., coding assistants).
- Development of best practices for defining "AI place" in learning (i.e., one fostering human growth and human critical thinking, with AI as an assistive vs supplantive technology).

B. Workforce Development

- A skilled workforce that can build and refine AI systems through activities ranging from data staging to full-scale model development.
- Just-in-time, embedded AI learning systems in work environments for ongoing upskilling.
- Advanced simulation environments to support accelerated acquisition of critical workforce skills, especially in areas vital to national security and economic leadership.
- A robust continuing education roadmap for upskilling and career advancement.

C. Higher Education and Integrity

- Engaging students in research with university faculty to train the next generation of AI developers.
- Policy and systems research to preserve academic integrity in the age of AI.

D. Instructional Integration

- Teaching students to apply AI in the development of high-quality applied solutions across every discipline.
- Using AI to enhance instruction and student engagement.
- Developing fundamental skills for AI fluency and ability to assess and evaluate appropriate use of AI.
- Preparing students to build custom AI-powered tools and applications.

7. Novel Mechanisms for Research Partnerships for Industry and Academia to Accelerate AI Development and Deployment

A. Develop consortia of multiple universities, national laboratories, and industry that directly addresses the challenges associated with developing novel shared intellectual property, streamlining and easing the approval burden and time taken to license technologies.

B. Establish national AI technology proving grounds, especially for cyber-physical systems and for going beyond leaderboards. This allows for rapid demonstration and real-world testing in controlled environments, creating unbiased benchmarks for safety, reliability, etc.; and assessing how AI models perform relative to pre-AI or hybrid alternatives on real-world tasks involving uncertainty, rare events, or critical decisions.

C. Invest in long-term sustainable cloud platforms to support the AI-powered tools and inference services produced by the consortium available to at low cost.