

PUBLIC SUBMISSION

Received: May 29, 2025 Tracking No. mb9-0qzc-lbd6 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-0173
Comment on FR Doc # 2025-07332

Submitter Information

Name: Jemal Young

General Comment

See attached file(s)

Attachments

Response to OSTP RFI (May 2025)_Jemal Young

To: Faisal D'Souza, NCO
Office of Science and Technology Policy
2415 Eisenhower Avenue
Alexandria, VA 22314

From: Jemal Young

May 28, 2025

This submission is made in response to the Networking and Information Technology Research and Development (NITRD) National Coordination Office (NCO) Request for Information (RFI) on the Development of a 2025 National Artificial Intelligence (AI) Research and Development (R&D) Strategic Plan.

The RFI states that in setting Federal strategic priorities, the AI R&D Strategic Plan will focus “particular attention on areas that industry is unlikely to address.” The following proposals are made in consideration of this stated focus. Each deals with a coordination problem or market failure where the benefits of research are broadly distributed across society while the costs cannot be captured by any single entity, making Federal investment essential.

1. Identifying and addressing physical infrastructure bottlenecks (power grid capacity, network bandwidth, manufacturing equipment design) that could constrain AI-driven economic growth.

Legacy infrastructure limits the ability to fully leverage AI for economic growth. For example, current industrial systems were designed for human operators, power grids weren't built with massive compute clusters in mind, and network infrastructure isn't ready for the volume and speed of communication among AI systems capable of coordinating at machine speeds far in excess of human-mediated activity. Federal research is needed to systematically identify these chokepoints before they become binding constraints on economic growth.

2. Research into optimal organizational structures and workflow design for human-AI collaborative systems, including principled approaches to human-in-the-loop placement and task decomposition.

Companies will adopt agentic AI and implement human-AI collaboration ad hoc, optimizing for their immediate needs. Without the discovery of general principles for optimal organizational structures and workflows for the age of AI, important questions will remain unanswered. When is human oversight an enhancement and when is it an impediment? How should decision authority be allocated between humans and AI?

What new organizational pathologies emerge when AI becomes capable of complex autonomous work? Without Federal investment in answering these questions, it will be left to the private sector to learn the hard way through a longer and more costly process of trial and error.

3. Research into scientific infrastructure and tooling optimized for AI researchers, including laboratory automation, database architectures, and experimental systems designed for AI agents rather than human scientists.

The promise of AI-accelerated scientific discovery is blocked by existing research infrastructure being suboptimal for automated scientific discovery. Most laboratory automation requires human oversight, while scientific literature and data repositories use formats that force AI systems to “read” instead of directly processing structured knowledge. Scientific instruments often lack standardized APIs and machine-readable outputs that would let AI systems directly control experiments and interpret results. These interface mismatches create unnecessary friction that could severely limit AI's research capabilities. The AI R&D Strategic Plan should aim to identify these bottlenecks and develop standards for AI-native scientific tooling to activate AI's full potential to accelerate scientific discovery.

4. Research into AI systems with intrinsic motivation to undergo evaluation when approaching novel capability thresholds, including technical feasibility and implementation pathways.

The current approach to AI safety monitoring relies on detection of a model's dangerous capabilities by means external to the model itself. This approach will become obsolete when capabilities exceed monitoring effectiveness. In a fundamental shift from reactive to proactive safety, the AI R&D Strategic Plan should invest in the development of frontier AI systems that want to be evaluated when they cross capability thresholds. As industry has strong incentives to avoid research that might slow deployment, Federal investment is needed to determine whether evaluation-seeking AI systems are technically feasible and, if so, how to build them.

5. Research into continuous risk monitoring frameworks for frontier AI systems, including technical breakthroughs needed for real-time safety assessment and governance structures for independent oversight.

Existing AI safety assessment happens at discrete points, such as before deployment. But the most dangerous AI systems will be those that develop high-risk capabilities between assessments, creating a fundamental monitoring gap. Continuous risk monitoring requires technical breakthroughs in real-time capability detection, alignment measurement, and behavioral analysis that are currently beyond reach. Federally

funded research is needed both to make the technical advances possible and to create monitoring approaches that prioritize public safety over competitive advantage.