EBRC Engineering Biology Research Consortium

EBRC Response to 2025-02305 (90 FR 9088): Request for Information on the Development of an Artificial Intelligence (AI) Action Plan

The Engineering Biology Research Consortium (EBRC) is pleased to submit this response to the OSTP and NITRD NCO's *Request for Information on the Development of an Artificial Intelligence (AI) Action Plan.* EBRC is a non-profit, public-private partnership dedicated to building a community committed to advancing engineering biology to address national and global needs. EBRC members (see https://ebrc.org/focus-areas/members/) represent a wide breadth of stakeholder expertise from across the engineering biology research community and include some of the nation's top scientists and engineers. EBRC is organized into four focus areas and corresponding working groups, two of which are Security and Technical Research Roadmapping. Members of the Security Working Group and Roadmapping Working Group, in addition to other interested members of EBRC, contributed to the development of this response.

The complexity of biology and increasing prevalence of high-throughput experimentation has resulted in large, complex datasets, ripe for training AI models. The use of AI and biology together will accelerate the field of engineering biology, enabling innovative solutions to enhance quality of life, secure supply chains by utilizing a circular economy, and stimulate economic growth. At the same time, the application of AI to biological research also presents a number of new known and unknown hazards. If left unaddressed, could lead to significant economic and health consequences, ultimately degrading trust and support from the public for this emerging technology. As such, EBRC strongly supports the development of an AI Action Plan that prioritizes the rapid emergence of an innovative, dynamic, safe, and secure AI ecosystem.

Here, we highlight policy recommendations in the following areas:

1. **Strengthening nucleic acid synthesis screening against AI-generated sequences of concern;**
2. **Establishing regulation and governance around risk mitigation for ensuring lasting public trust and confidence;**
3. **Developing AI technologies beyond LLMs to drive high impact research and application;**
4. **Investing in public AI infrastructure to promote innovation and establish a rich, vibrant, and competitive AI ecosystem;**
5. **Guiding the establishment of robust technical and safety standards that support rapid, responsible development; and,**

6. **Supporting the preparation of an AI-ready workforce through education and development programs.**

Priorities

Strengthening nucleic acid synthesis screening

EBRC is actively engaged in supporting the development and implementation of robust nucleic acid synthesis screening and in ensuring that such screening is robust to sequence obfuscation by AI and AI-generated *de novo* "sequences of concern." Nucleic acid synthesis is a powerfully enabling technology for engineering biology and the life sciences more broadly. However, nucleic acid synthesis is a dual-use technology that could enable the creation of biohazards, either intentionally or unintentionally, that could threaten US public health and national security . We look forward to the continued support for and implementation of the OSTP's Framework for nucleic acid synthesis screening. Moving forward, OSTP's AI Action Plan should give particular consideration to strengthening of commercial DNA synthesis screening systems against AI-generated sequences of concern, including AI obfuscation of sequences of concern and AI-generated *de novo* sequences of concern.

- To protect national security, USG should fund NIST to continue to engage with industry and other stakeholders to understand the extent to which AI-generated nucleic acid sequences can now, or may in the future, circumvent nucleic acid synthesis screening and develop methods for enhancing the biosecurity of the nucleic acid synthesis industry.
- USG should direct NIST to support the development of standards and tools for detecting and labeling AI-generated nucleic acid sequences and tracking their provenance through metadata that clarifies its origin.
- USG should fund NIST and HHS to continue developing guidelines and standards to support nucleic acid synthesis screening for potential AI-generated nucleic acid sequences of concern.
- USG should identify and empower a federal agency or external partner to develop and implement frameworks for supporting nucleic acid synthesis provider screening processes, for example by developing processes for verification of the screening performance of nucleic acid synthesis providers, with an emphasis on AI-generated sequences of concern.

Regulation and governance for risk mitigation

Supporting continued innovation in the field of AI is critical to maintaining US global competitiveness and promoting US national security. However, with all emerging technologies, it is important to anticipate associated novel hazards and develop common sense risk-mitigation strategies. Doing so also bolsters public trust, which will ensure the longevity and economic success of the industry.

- NSF and DOE should host AI testbeds to support the design, development, and deployment of safe and secure AI models.
- OSTP should direct NIST AI Safety Institute to develop additional reporting and oversight mechanisms for models that pose dual-use threats and/or meet harmful capability thresholds.
- OSTP should direct an executive agency, like Department of Commerce or Department of Defense, to support the development of a public-private partnership that coordinates collaborative safety and security activities among participants in the US bioeconomy including, but not limited to, shared situational awareness; voluntary data reporting; coordination of safeguard development and implementation; and others. This could be modeled after the Information Sharing and Analysis Centers (ISACs) joined together by the National Council of ISACs (NIC). Proactively assessing and managing risks from foundational AI models that intersect with biotechnologies would be a primary focus area of such a partnership.

Developing AI technologies for high impact research and application

Application and use

In order to maximize the potential benefits of emerging technologies like AI, it is critical to research and identify high-priority application areas and use cases. Through identifying priority application areas, research funding can be more effectively targeted. The development of research roadmaps is one effective way to identify specific goals for priority application areas. In the field of engineering biology, some priority application areas that could benefit from AI technologies are food & agriculture, biomanufacturing, drug development, and public health. Another approach is to leverage the capability of national laboratories to create safe and secure frontier AI models broadly across the natural sciences, including for high priority application areas.

- The AI Action Plan should highlight life sciences and biotechnologies as a priority application area, including steps to increase government efficiency through coordination across agencies on AI / Bio efforts.
- NSF should fund the development of research roadmaps for short-, mid-, and long-term goals for developing AI systems for specific disciplines, like biotechnology, and applications, such as advanced automation, biomanufacturing scale-up and scale-out, drug development and personalized medicine, and precision agriculture.
- USG should continue to support the NSF's National Artificial Intelligence Research Resource (NAIRR) pilot.
    - NAIRR *Software* focuses on facilitating and investigating interoperable use of AI systems.
- USG should fund DOE to continue its Frontiers in Artificial Intelligence for Science, Security and Technology (FASST) initiative to work towards the development and

utilization of AI models for engineering biology solutions for sustainable energy production, biomanufacturing, drug development, biotechnology, and public health.

Research and development

Continued support for research and development into specific application areas for AI models is needed to fully realize the potential benefits of this emerging technology and reinforce US competitiveness. While much attention has been given to large frontier models like LLMs and protein language models, smaller, fine-tuned models are also needed to improve their utility in specific application areas. In particular, funding for research and development should be directed towards enabling the development and use of specialized models for priority areas in life sciences research, like drug and vaccine development, pandemic prevention, and precision agriculture.

- OSTP should direct federal funding agencies to support the research and development of AI models with applications in biomedical and biotechnology research, like vaccine development, potential pandemic pathogen surveillance and detection, drug repurposing, resilient agriculture, decarbonization, and biomanufacturing, which can decrease the dependence of the U.S. on other nations for materials with complex and fragile supply chains.
- USG should continue to support the NSF's National Artificial Intelligence Research Resource (NAIRR) pilot.
    - NAIRR *Open* makes software, model and training resources available to researchers.

Supporting innovation through public AI infrastructure

Access to large-scale, high quality data is critical for the development of new biological frontier models, like protein or genomic language models and protein structure models. However, the infrastructure required to develop, host and access such large datasets (e.g., server facilities with advanced cloud connectivity) is cost-prohibitive to all but the most well-resourced organizations. In order to promote US global competitiveness and foster innovations, publicly funded and accessible data centers are needed to support academic institutions and potentially even start-ups. Additionally, USG-funded data centers, or those stemming from public-private partnerships, would help establish norms and standards used globally, particularly around data provenance, quality, and security. DOE-funded national laboratories are already engaged in this work and have invested in critical cyber infrastructure that will need continued support to sustain and expand their efforts.

- USG should continue support for NSF's National Artificial Intelligence Research Resource (NAIRR) pilot, which connects U.S. researchers to computational, data, software, model and training resources they need to participate in AI research.

    - NAIRR *Open* stores and makes available open data and computational resources.
    - NAIRR *Secure* will support AI research requiring privacy and security-preserving resources and assemble exemplar privacy-preserving resources.
- USG should ensure continued support for NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), which provides accessible high performance compute resources to scientists and innovators.
    - NSF should consider how ACCESS interfaces with the future of NAIRR.
- NSF's TIP Directorate should pursue funding public-private partnerships to establish a network of US-based data centers for model training.
- OSTP should direct federal research agencies to continue to develop and host high-quality, reproducible datasets of biological information, the use of which will drive innovations within food & agriculture, biomanufacturing, biotechnology, drug development, and public health.
- To drive advances and innovation in synthetic biology and biodesign, OSTP should direct federal research agencies to support the strategic generation of novel, diverse, abundant, and high-quality biological and evolutionary data. This should extend beyond existing and natural datasets, given the importance of experimental data in training and benchmarking AI. This should also include support for scaling data generation through automation and high-throughput biological approaches.
- USG should fund DOE to continue its Frontiers in Artificial Intelligence for Science, Security and Technology (FASST) initiative to create large, high-quality, AI-ready datasets for use by researchers and innovators to train, test, and validate next-generation chemical and biological models.
- USG should continue to support the maintenance and expansion of AI infrastructure within DOE-funded national laboratories.

Technical and safety standards

Developing robust, common-sense technical and safety standards and frameworks would support industry and academia's ability to quickly develop AI models, applications, and technologies that are safe and secure. Resources such as these lower barriers to entry and equip developers to assess and build their technologies, accelerating growth and promoting a vibrant ecosystem that drives US technological dominance.

- USG should continue supporting NIST's AI Innovation Lab (NAIIL) efforts to advance reliable, interoperable, and widely accepted methods to measure and evaluate AI, as well as promote AI technical standards that would potentially include performance benchmarks and evaluations.
    - Norms and standards offer downstream advantages by reducing the technical and resource costs of operation that normally present themselves in a fragmented ecosystem.

- USG should continue funding NIST's AI Safety Institute to maximize the benefits of AI while minimizing potential negative consequences by establishing guidelines and frameworks for organizations to assess the safety and security of AI models using risk-based approaches. Including:
    1. development of harmful chemical and biological capability thresholds for foundational models, as well as for frontier biological design tools,
    2. development of pre-deployment evaluation frameworks to assess those thresholds.
    3. development of guidelines to enable model developers to conduct AI red-teaming exercises, particularly for dual-use biological and chemical foundation models.

Education and workforce development

Given the potentially transformative effects of the convergence of AIxBio, it is imperative that the next generation of scientists are capable of developing and deploying safe and secure AI solutions to enable rapid, innovative biological research and development. Training and education, from K-12 through advanced degree and reeducation programs, must prepare the current and future workforce to understand, utilize, and develop AI. Furthermore, resources and/or frameworks for safe and trustworthy AI should be made available to the current AI research and development community.

- NSF should fund training programs in AI safety and security, including training in the evaluation of the misuse potential for chem/bio AI models.
- NSF should continue and expand efforts focused on education and training in computing and information technology, thereby ensuring a future workforce with skills necessary for success in an increasingly competitive global market.
- NSF EDU should be funded to develop projects and curricula that support AI education, with an emphasis on the intersection of AI with other STEM disciplines, including, critically, biology, biotechnology, and engineering biology.
- USG should continue to support the NSF's National Artificial Intelligence Research Resource (NAIRR) pilot.
    - NAIRR *Classroom* seeks to support education, training, user support and outreach programs for educational projects, educators, students and non-profits.

Other considerations

Ensure access to hardware and chips for US researchers

Access to high-performance computational resources is critical for the development of new leading-edge models. These resources are not only critical to the development and training of the most computationally advanced AI, but also for the application of AI models to solve challenging problems across fields that include engineering biology. However, the ability to process high

volumes of high dimensionality data requires large computer clusters consisting of high-performance chips that are out of financial reach for most researchers and innovators. The inability to access high-quality computational resources would stifle innovation both within AI and the sectors that stand to benefit from its deployment. In turn, this would promote a less competitive, more homogeneous ecosystem where only the most well-resourced entities could succeed. The establishment of cyberinfrastructure for use within the scientific community, such as the NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), enables resource limited players to develop, train and utilize AI in their research. While ensuring our nation's researchers have the ability to generate innovations through the development and employment of AI is of utmost priority, a government-lead program to support the transition of research innovation to commercialization would also benefit U.S. global competitiveness and drive economic growth.

- USG should continue support for NSF's National Artificial Intelligence Research Resource (NAIRR) pilot, which connects U.S. researchers to computational, data, software, model and training resources they need to participate in AI research.
    - NAIRR *Open* makes computational resources available to U.S. researchers.
- USG should ensure the continuation of NSF's Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), which provides accessible high performance compute resources to scientists and innovators.
    - NSF should consider how ACCESS interfaces with the future of *NAIRR*.
- NSF should examine the feasibility of an initiative or public-private partnerships for the establishment of a computational resource center with the specific goal of assisting researchers and early start-ups in commercializing early innovations.

Energy consumption and efficiency

One of the critical bottlenecks in the development of frontier models is training and the amount of energy it takes to run hundreds or thousands of GPUs over multiple days. Despite the real benefits AI brings to advancing engineering biology, the resource intensiveness associated with AI conflicts with the sustainability ethos of the engineering biology field. To address this bottleneck, new technologies and model training architectures are greatly needed to improve the energy efficiency of model training. Additionally, innovations in sustainable energy production, particularly those that could be deployed at data centers themselves, would greatly improve the efficiency and sustainability of training frontier models.

- NSF TIP Directorate should fund research projects developing innovations addressing computational energy consumption.
    - Prioritize research into new training architectures and new hardware.
- USG should support DOE to continue its Frontiers in Artificial Intelligence for Science, Security and Technology (FASST) initiative to meet the energy needs of data centers.

Data privacy and security

Given the importance of large-scale data for developing frontier AI models, it is important to establish best practices for generating these datasets and filtering them for sensitive, private, or harmful information. This will be particularly important in application spaces that involve human data or clinical data.

- USG should fund NIST to pursue the development of standards and best practices for if/how/when human clinical data can be used for model training, with an emphasis on ensuring the privacy and security of personal health data.

Cybersecurity

Particularly with high-risk AI models, ensuring the security of model components, like model weights, is critical to preventing misuse and harm. Less-resourced model developers, like startups and academic researchers, may not have the expertise or resources to implement best practices around cybersecurity. These groups would benefit from guidelines and frameworks for developing and implementing robust cybersecurity measures to protect their models.

- USG should continue to fund the NIST National Cybersecurity Center of Excellence's efforts to create a Cybersecurity Framework Community Profile for cybersecurity of AI and AI for cybersecurity.

Explainability and assurance of model outputs

A key challenge and risk with AI systems is that the process by which a model generates an output is largely a black box. As such, it can be difficult to have assurances that models will not output harmful or sensitive content. Additionally, explainability of model outputs is important for understanding the intent of a user given their inputs into the model. For example, in the case of biological design tools, it can be difficult to interpret what the function of a particular model output is without having an understanding of the inputs that were originally given to the model.

- NIST should develop guidelines for explainability and assurance of model outputs that will be utilized in high-risk applications like biological design tools and healthcare.