**Comments of the Software & Information Industry Association (SIIA)**

**U.S. Department of Energy Request for Information on Frontiers in AI for Science, Security, and Technology (FASST) Initiative**

**November 11, 2024**

The Software and Information Industry Association (SIIA) appreciates the opportunity to provide comments in response to the request for information on the Frontiers in AI for Science, Security, and Technology (FASST) initiative issued by the Department of Energy (DOE). We believe the FASST initiative is among the most significant efforts undertaken by the U.S. government to realize the potential of AI and advance U.S. economic and national security. We commend DOE for seeking public input at this stage.

SIIA is the principal trade association for companies in the business of information. Our members include nearly 400 companies reflecting the broad and diverse landscape of digital content providers and users in academic and scientific publishing, education technology, and financial information, along with creators of software and platforms used by millions worldwide, and companies specializing in data analytics and information services.

SIIA has been an avid supporter of FASST and was an early endorser of the Department of Energy AI Act of 2024 (H.R. 9671, S. 4664), a bill in Congress that would formally authorize FASST and support needed for DOE to undertake this visionary program.[1] DOE and the 17 national laboratories have a critical role in advancing AI-driven scientific discovery and applications, owing in part to the tremendous potential value of DOE's datasets and advanced computational resources that DOE has developed.[2] By fostering long-term strategic planning

---

[1] SIIA, "SIIA Supports the DOE AI Act" (Jul. 16, 2024) https://www.siia.net/siia-supports-the-doe-act/; SIIA, "SIIA Supports AI Legislation on Incident Reporting and Department of Energy R&D" (Oct. 1, 2024) https://www.siia.net/siia-supports-ai-legislation-on-incident-reporting-and-department-of-energy-r-praises-house-science-committee-for-sending-legislation-to-the-floor/

[2] See, e.g., Jonathan Carter, et. al., *Advanced Research Directions on AI for Science, Energy, and Security: Report on DOE Summer 2022 Workshops* (May 2023), https://www.anl.gov/sites/www/files/2023-06/AI4SESReport-2023-v6.pdf (hereinafter "DOE, *AI for Science*"); Testimony of Dr. Georgia Tourassi, Associate Laboratory Director for Computing and Computational Sciences, Oak Ridge National Laboratory, before the House Science, Space, and Technology Committee (Feb. 6, 2024), https://republicans-science.house.gov/_cache/files/c/0/c03e8d60-faa8-4b06-97c7-f60031a6da5b/853C57C2116017E07EA69E1BF543E23B.2024-02-06-tourassi-testimony.pdf; Testimony of Rick Stevens, Associate Laboratory Director for Computing, Environment, and Life Sciences, Argonne

and interagency collaboration that are essential to maintaining U.S. leadership in AI innovation as well as national security, we believe FASST will be transformational. We believe that achieving that goal will require close collaboration between DOE and industry.

Our comments focus on how DOE can, through FASST, develop a nationwide, comprehensive data ecosystem to support advanced AI research. We believe this work can serve as a model for the government and strengthen U.S. leadership in AI-driven research and applications.

**1. Data**

DOE and the 17 national labs possess extraordinarily rich data resources that have the potential to yield world-changing insights and applications through the FASST initiative. In response to this portion of the RFI, we provide two recommendations that we believe will help DOE to unlock this potential.

*Develop data governance practices that foster interoperability.*

We recommend that FASST strive to develop data governance practices and standards that foster interoperability. The Office of Management and Budget (OMB) defines interoperability as "the ability of two or more systems, products, or components to exchange information and use the information that has been exchanged, including to operate effectively together. This includes ensuring that open and standard data formats and application programming interfaces (APIs) are used so that foundational components can be used, including to build for new use cases, without obscure proprietary technologies or licensing."[3]

Though OMB guidance focuses on interoperability as relevant to promoting competition and avoiding vendor lock-in, it is critical to enable more powerful AI-driven research. As the *AI for Science, Energy, and Security* report explains, applying findable, accessible, interoperable, and reusable (FAIR) data standards is essential for training and across "every AI application area" including "the training process and the models themselves."[4] Combining open data sets with government data sets, and data generated from across the national labs, is a significant barrier to creating new foundation models and pursuing new lines of inquiry. In short, it is important

---

National Laboratory, before the Senate Energy & Natural Resources Committee (Sept. 7, 2023), https://www.energy.senate.gov/services/files/CF8309D8-C0A1-40C7-944F-CF71EF523FF8

[3] OMB, M-24-18: Advancing the Responsible Acquisition of Artificial Intelligence in Government (Sept. 24, 2023) https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf

[4] DOE, *AI for Science* at 75. Argonne National Laboratory's Rick Stevens, a co-author on the *AI for Science* report, noted that realizing the potential of FASST would require "on the order of 2,000 more staff" to develop and evaluate new models and reorganize scientific data into workable formats. Johnny Miri, "'AI for Science' Initiative at DOE Gains Traction," *AIP FYI* (Jun. 6, 2024) https://ww2.aip.org/fyi/ai-for-science-initiative-at-doe-gains-traction



that DOE create and implement standards that support interoperability to break down barriers across data silos.

*Incorporate best-in-class privacy enhancing technologies (PETs).*

We further recommend that DOE focus on ways that privacy enhancing technologies (PETs) can enable AI and machine learning research that provide the appropriate level of privacy and security required for various datasets.

To deliver the best outcomes, AI and machine learning capabilities need to be trained, enriched, and leveraged over a broad, diverse range of data sources. When a machine learning model is trained on new and disparate datasets, it becomes smarter over time, resulting in increasingly accurate and valuable insights. However, models encode the data over which they were trained, introducing notable privacy and security risks if data sources contain PII, intellectual property, or other forms of sensitive or restricted data. This can create potential vulnerabilities, and particularly in the context of sensitive or classified data that may be involved in specific research endeavors, represents a potential national security risk.[5]

As is often the case, the risks associated with technological advances such as those we're currently seeing in the AI arena can be offset with other breakthroughs in technology. PETs are a family of technologies that protect data while it is being used or processed. They not only reduce those risks, but also expand the way organizations such as DOE can leverage data to unlock value for a national AI capability.

PETs uniquely address the challenges of securely leveraging disparate data sources by preserving the privacy of data while it is being used. For AI use cases, they allow users to securely train and evaluate ML models using data sources across silos and boundaries, including cross-jurisdictional, third-party, and publicly available datasets. By protecting AI models and workflows during processing – i.e., "data in use" – PETs can enable secure AI capabilities that enhance scientific research, discovery, and implementation. In addition to enabling net-new data usage, PETs also help ensure sensitive assets, including ML models trained over regulated data sources, remain protected at all points in the processing lifecycle. This limits the increased risk presented by even the most complex threats within the AI landscape such as data spoofing, model poisoning, and adversarial ML.

In addition to model vulnerabilities, DOE should also consider that AI/ML models are data hungry. The need to aggregate data in a centralized location for model training is often a deal-breaker for utilization, especially when the data in question is sensitive in nature. PETs eliminate

[5] Vulnerabilities in ML models typically lead to two macro categories of attack vectors: model inversion and model spoofing. Model inversion attacks involve targeting the model itself to reverse engineer back to the data over which it was trained — data that is likely sensitive and therefore valuable to an attacker. Model spoofing, on the other hand, represents a form of adversarial machine learning wherein an attacker attempts to deceive the model by manipulating the input data in such a manner that the model makes incorrect decisions aligned with the attacker's intentions.



the need to replicate or pool data, allowing data owners to retain positive control of their assets and limit the risk of misuse and unintended exposure. These solutions give users flexibility, enabling data value extraction while protecting the interests of all stakeholders.

PETs deliver value in a variety of use cases across commercial and public sector scenarios, and are in use today, enabling secure data usage in ways that were previously impossible. Organizations can securely enrich existing ML models by expanding the pool of available data sources for training and deployments options. By uniquely protecting the AI/ML model, PETs allow organizations to leverage models trained on sensitive data over new datasets that might otherwise be unusable. Securing the model mitigates risks, and organizations can utilize AI capabilities over a broader scope of datasets, ultimately delivering deeper, richer insights. Encrypted models can be leveraged across security, third-party, and organizational boundaries, even when using highly sensitive or proprietary models, allowing organizations to securely and privately derive insights and improve outcomes.

**2. Compute**

We recommend that DOE prioritize the development of a comprehensive data ecosystem built on a hybrid cloud environment that provides interoperability across cloud service providers and systems.

As an initial matter, there exist significant institutional and regulatory barriers to realizing the potential of FASST and unlocking the value of DOE's vast datasets through AI technologies. As detailed in the *AI for Science, Energy, and Security* report, barriers to developing a comprehensive data ecosystem include security considerations, disparate data governance and management practices across DOE facilities, lack of protocol and standardization tools to promote data sharing, ad-hoc or domain-specific practices, the sheer volume and variety of DOE initiatives, and the resources required to implement FAIR data standards.[6] In addition to these, the traditional means of procuring cloud services leads to adoption of a variety of environments that generally are not interoperable, both within particular facilities and across DOE and the 17 national labs. Lock-in with single vendors can lead to security and resiliency challenges and disincentives innovation that will further the DOE mission. Moreover, given the rapid advances in commercial cloud infrastructure, DOE's limited use of commercial cloud resources constrains its ability to use the most advanced AI tools for research and to harness the compute power available from commercial vendors. There is currently a lack of adequate cloud adoption, with many organizations choosing on-premises workloads. Lastly is the issue of scale. We believe it necessary that DOE work with industry to build capacity at the scale necessary to undertake its ambitious program. As one official has described it, the FASST models could "completely consume" DOE's supercomputers.[7]

---

[6] DOE, *AI for Science*, at 127-128.

[7] According to *AIP FYI,* "DOE would need to build the capacity to support 'something like 200,000' researchers using the AI models each day once they are built, based on the current staff size of the



To address these challenges, we propose an approach with the following key features:

- Multi-cloud. We recommend a multi-cloud approach to support resiliency, security, and take advantage of innovative solutions.[8] By "multi-cloud," we mean using multiple cloud service providers, such as commercial cloud solutions based on mission and engineering requirements, allowing DOE to develop what they deem best for different components of the FASST project. The Department of Defense (DoD) is among those in the federal government that has realized the importance of multi-cloud and, as it finalizes its new cloud strategy, could provide insight to operationalize a similar approach across DOE.[9]

- Hybrid-cloud. We further recommend that DOE pursue a hybrid-cloud approach, one that takes advantage of the compute power and software resources available only in commercial cloud environments, seeking to expand capacity of classified environments hosted on commercial cloud for classified processing. For example, many if not all AI foundation models can be used only in commercial cloud environments owing to government cloud restrictions. Moreover, it will be impossible for DOE to attain the scale contemplated in the FASST initiative without making use of commercial compute.[10]

---

national lab system and its user base," quoting Argonne's Rick Stevens. Johnny Miri, "'AI for Science' Initiative at DOE Gains Traction," *AIP FYI* (Jun. 6, 2024) https://ww2.aip.org/fyi/ai-for-science-initiative-at-doe-gains-traction

[8] SIIA, et al., Letter to Executive Branch Leaders on reducing federal cybersecurity risk (Jun. 12, 2024), https://www.siia.net/wp-content/uploads/2024/06/Industry-Cybersecurity-Letter-to-Executive-Branch-June-12-2024-1.pdf

[9] *See, e.g.*, Grace Dille, "DoD's Lamb: Multi-Cloud Strategy, Software Directive Coming Soon," *MeriTalk* (Aug. 6, 2024), https://www.meritalk.com/articles/dods-lamb-multi-cloud-strategy-software-directive-coming-soon/; "DISA looks to offer multi-cloud capabilities at the tactical edge," *DefenseScoop* (May 21, 2024), https://defensescoop.com/video/disa-looks-to-offer-multi-cloud-capabilities-at-the-tactical-edge/; Calvin Hennick, "Defense Agencies Turn to Multicloud Strategy," *FedTech* (May 22, 2024), https://fedtechmagazine.com/article/2024/05/defense-agencies-turn-multicloud-strategy; John B. Sherman, DoD CIO, Memorandum on DoD Joint Warfighting Cloud Capability (Jul. 31, 2023), https://dodcio.defense.gov/Portals/0/Documents/Library/NextStepRationalizeCloudUse.pdf

[10] A massive increase in computational resources is also required for AI-enabled workflow optimization. As stated in the *AI for Science* report: "The AI models described throughout this report will require extensive computational resources for training and execution, with the potential for inverse design capabilities that could themselves be used to propose both improvements in resource use and new designs for resources—from instruments to supercomputers. These designs could, in turn, drive AI-enabled automated design and manufacturing to orchestrate their construction, operation, and use. For the scale and uniqueness of DOE mission areas, the realization of these advances will require building infrastructure to support the embedding of AI in workflow systems, incorporating performance and results data to continuously self-train, and advancing workflow technology to enable further breakthroughs in the use of AI for DOE mission areas." DOE, *AI for Science*, at



Commercial cloud environments and the AI tools available in those are likely suitable for a range of scientific research, as has been demonstrated by advanced genetics work undertaken by DeepMind.[11] Commercial clouds can provide security and reliability; some providers offer commercial cloud solutions with the capacity to process sensitive and classified information.[12] We recommend DOE conduct further exploration into processing classified data with input and assistance from providers of both commercial cloud solutions and PETs.

- Interoperability. Critical to the success of FASST – and to a hybrid, multi-cloud approach – is ensuring that systems are interoperable with one another. Whether through APIs or other measures,[13] we recommend that DOE prioritize the development of interoperability and open standards to maximize the ability to work across the DOE enterprise as well as with key academic and industry partners. Further, coordinated data sharing policies across DOE components and labs is essential in several sectors for research and development as well as an improved end-user experience.

Pursuing the type of approach outlined here has two additional advantages that go beyond the FASST initiative. First, developing a comprehensive data – and compute – environment in FASST will serve as a model that can be adopted across the federal government. Indeed, we believe that the government should adopt a comprehensive data and cloud strategy that leverages commercial capabilities and enables intra- and inter-agency uses as needed.[14] This is a challenge that has impeded adoption of AI, and the insights that AI can glean for a range of government services and operations, not only for DoD but across the federal government.

Second, an integrated, comprehensive approach to data and compute will advance U.S. leadership in AI. China has already seized on this approach, as reflected in China's "Eastern Data

[11] See also, e.g., National Academies of Sciences, Engineering, and Medicine, et al., *Charting a Path in a Shifting Technical and Geopolitical Landscape: Post-Exascale Computing for the National Nuclear Security Administration* (2023) at 58-62 https://nap.nationalacademies.org/catalog/26916/charting-a-path-in-a-shifting-technical-and-geopolitical-landscape

[12] See Testimony of Rick Stevens, Associate Laboratory Director for Computing, Environment, and Life Sciences, Argonne National Laboratory, before the Senate Energy & Natural Resources Committee (Sept. 7, 2023), https://www.energy.senate.gov/services/files/CF8309D8-C0A1-40C7-944F-CF71EF523FF8

[13] OMB, M-24-18: Advancing the Responsible Acquisition of Artificial Intelligence in Government (Sept. 24, 2023), https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf (requiring agencies to prioritize interoperability "to the greatest extent practicable").

[14] James Andrew Lewis, "Accelerating Federal Cloud Adoption for Modernization and Security," CSIS (Jul. 28 2023), https://www.csis.org/analysis/accelerating-federal-cloud-adoption-modernization-and-security



and Western Computing" initiative.[15] The initiative is designed in part to promote energy efficiency by relocating data centers to areas with low populations and rich resources. It is also designed to enable the integration of compute resources and interoperability of AI models. The United States lacks a comprehensive strategy to harness the potential of optimizing the insight from across data silos and AI models and one that allows multiple stakeholders --- academic and private sector researchers – to engage with one another to improve research and applications. It means that the United States risks falling behind as competitors like China pursue holistic approaches to AI.

**3. Models**

We support DOE's interest in developing science-oriented foundation models. We recognize this as a strategic, critical need that has been historically underdeveloped by the market, and an appropriate area for DOE to help incentivize development.

**4. Applications**

We recommend that the FASST initiative include efforts to deploy AI technologies to optimize energy efficiency; developing and optimizing new energy sources, such as nuclear and fusion; advancing electrification and grid transformation; and improving grid resilience. Notwithstanding the many areas where FASST can contribute to developing scientific research, DOE is unique in its ability to convene experts to solve the nation's greatest energy challenges.[16] AI tools are uniquely suited to provide ," guidance to address the tremendous energy required to continue to fuel AI innovation and applications.[17] We envision AI-driven energy consumption

[15] Ning Zhang, et al., "The 'Eastern Data and Western Computing' Initiative in China contributes to its net-zero target," *Science Direct* (Aug. 20, 2024), https://www.sciencedirect.com/science/article/pii/S2095809924005058; Huawei, "Eastern Data and Western Computing: Building New Computing-first Networks," *Huawei Tech* (Mar. 2022), https://www.huawei.com/en/huaweitech/publication/202202/eastern-data-western-computing-network

[16] See, e.g., Dept. of Energy, *AI for Energy: Opportunities for a Modern Grid and Clean Energy Economy* (April 2024), https://www.energy.gov/sites/default/files/2024-04/AI%20EO%20Report%20Section%205.2g%28i%29_043024.pdf

[17] See, e.g., Alex de Vries, "The Growing Energy Footprint of Artificial Intelligence," *Joule* (Oct. 18, 2023), https://www.cell.com/joule/fulltext/S2542-4351(23)00365-3; Olivier Cognet, "Use Smart Digital Solutions forecasting Tools to Manage Power Grids," *Schneider Electric Blog* (Jul. 15, 2020), https://blog.se.com/energy-management-energy-efficiency/2020/07/15/use-smart-digital-solutions-forecasting-tools-manage-power-grids/; Vida Rozite, et al., "Why AI and Energy Are the New Power Couple," International Energy Agency (Nov. 2, 2023), https://www.iea.org/commentaries/why-ai-and-energy-are-the-new-power-couple



to continue to grow, and developing AI tools to create grid efficiencies and develop new energy sources is an important and underdeveloped application.

In addition, we recommend that DOE consider launching grand challenges that focus on "strategic and critical application spaces that would otherwise be underinvested."[18]

**5. Workforce**

DOE recognizes the urgent need to grow the AI workforce and the AI-educated science workforce, reflected in the dozens AI workforce training programs already in place.[19] SIIA supports additional efforts to address the shortage of highly skilled AI talent needed to maintain U.S. competitiveness. We believe the private sector should be a resource to help address this challenge. We would encourage DOE to engage with the private sector to explore the following:

- Investing in AI and STEM education at K-12 and at the university level;
- Developing training programs for government employees;
- Creating a fellowship program to enable experts from the private sector to work at DOE and the national labs for a defined term, modeled on the Intergovernmental Personal Act Mobility Program; and
- Creating a fellowship program to enable researchers from DOE to embed with private sector companies for a defined term to develop expertise in critical areas of AI development.

We also support several bills under consideration by Congress that will assist DOE and the federal government at large in developing the workforce of the future, including:

- H.R. 9671/S. 4664, the Department of Energy (DOE) AI Act of 2024. Discussed above, this bill will provide DOE with authorities for new STEM education and workforce development programs.
- H.R. 9211, the LIFT AI Act. This bill is intended to increase investment in AI literacy, preparing students for success in a changing workforce.
- S. 4394/H.R. 9402, the National Science Foundation Artificial Intelligence Act. By incentivizing public-private partnerships, this bill will empower students with access to

[18] Dept. of Energy, FASST Fact Sheet, https://www.energy.gov/sites/default/files/2024-07/FASST%20Handout%20%281%29_0.pdf

[19] Dept. of Energy, Office of Critical and Emerging Technologies, "Supercharging America's AI Workforce," https://www.energy.gov/cet/supercharging-americas-ai-workforce



cutting-edge AI technologies, ensuring the next generation is well-prepared for the future workforce.

- H.R. 9215, the Workforce for AI Trust Act. This bill builds on the National AI Initiative Act of 2020 to promote developing experts to build trustworthy AI systems.

- H.R. 9403, the Expanding AI Voices Act. This bill will advance AI innovation by fostering partnerships and investing in higher education and other institutions to expand AI capacity in populations historically underrepresented in STEM.

**6. Governance**

Given the scale and vision of FASST, we recommend that DOE work in close partnership with the private sector and academic institutions to develop and implement the initiatives. The private sector's development of PETs, cloud infrastructure, foundation models, and other essential elements of the AI stack are necessary to achieve the FASST mission.

Flexibility in procurement is a necessary aspect of the public-private partnership needed for FASST. The DHS Procurement Innovation Lab and NASA's Acquisition Innovation Launchpad are two examples of efforts within the federal government focused on agile procurement. We believe that by opting for a multi-vendor approach (as reflected in our earlier points about interoperability across cloud providers and systems) and promoting interoperability of data and systems, DOE can avoid challenges of lock-in, vendor preferencing, and limitations of legacy systems. This will help to ensure a competitive playing field and mitigate against conflicts of interest. DOE should also explore partnering with the Small Business Administration, the NSF Office of Small and Disadvantaged Business Utilization, and similar federal agencies to ensure smaller businesses with cutting-edge technology have an opportunity to participate in building out the FASST initiative.

Beyond procurement, however, DOE should leverage the exceptional talent and expertise in industry, academia, and civil society through a streamlined group of advisory councils. NIST's experience with the AI Safety Institute Consortium should provide lessons learned about how to convene a multitude organizations and individuals with expertise relevant to the government mission. We recommend that DOE take care to create an advisory council structure that is designed to the mission, which may counsel in favor of a smaller number of participants representative of industry, academia, and civil society. The Defense Business Board and the President's Council of Advisors on Science and Technology are two that seem to strike this balance. Trade associations like SIIA can be useful participants because they represent a large number of companies.

We further recommend that DOE institute a robust interagency process to leverage expertise from other agencies (including NIST, the NSF, DoD, and the intelligence community) and to provide guidance to assist transformation across the government. As part of this, DOE and the national labs should work with NIST in developing open standards and technical guidance both for interoperability and for developing a hybrid, multi-cloud environment.



*  *  *

SIIA appreciates the opportunity to provide our views on the FASST initiative. We are enthusiastic about this initiative and its potential to have a significant impact on AI science, U.S. leadership in AI, and national security. We look forward to collaborating with DOE as this work progresses. For inquiries, please reach out to Paul Lekas, SIIA's Head of Global Public Policy and Government Affairs, at plekas@siia.net.