RFI Response

Artificial, Inc. and Prof. Jason Hattrick-Simpers of the University of Toronto (also part of Acceleration Consortium) are pleased to provide our response to the Department of Energy's (DOE's) Request for Information (RFI) on the Frontiers in AI for Science, Security, and Technology (FASST) initiative. Our experience in AI-driven scientific workflows, particularly in areas like new materials discovery, drug development, and self-driving labs, aligns with the DOE's mission to capitalize on AI for advancing scientific discovery and applied energy solutions.

Respondent Information

The Department of Energy (DOE) requires the following respondent information in RFI submissions. Below is the required information for Artificial, Inc.'s response to the Frontiers in AI for Science, Security, and Technology (FASST) initiative:

- **RFI Title**: Notice of Request for Information (RFI) on Frontiers in AI for Science, Security, and Technology (FASST) Initiative
- **Principal Points of Contact**:

  - Dr. Yao Fehlis
    - Phone: 1-8324991256
    - Email: yao@artificial.com
  - Prof. Jason Hattrick-Simpers
    - Email: jason.hattrick.simpers@utoronto.ca

- **Institution or Organization Affiliation**:

  - Artificial, Inc., 380 Portage Ave, Palo Alto, CA, United States 94306
  - University of Toronto, 27 King's College Cir, Toronto, ON M5S 1A1, Canada
  - Acceleration Consortium, 700 University Ave suite 1001, Toronto, ON M5G 1X6, Canada

- **Clear Indication of the Specific Questions**:
  - Data: (c) and (d)
  - Models: (a) and (b)
  - AI Applications: (a) and (b)

This structured format ensures easy access to vital contact and organizational details as required by the DOE.

Data

(c) Are there partner organizations with relevant scientific or energy-related data that DOE should work with?

Organizations such as **National Institute of Standards and Technology (NIST), the Acceleration Consortium (AC),** and the **University of Toronto** play crucial roles in providing high-quality datasets and research expertise in fields like materials science and energy technology. NIST is renowned for its work on measurement standards and advanced technology development, making it an essential partner for AI research and development. Similarly, the University of Toronto is a leader in AI and machine learning research, with extensive work in computational science and data analytics applicable to energy and technology sectors. The AC is a world leading consortium, based in Toronto, that specializes in the development and implementation of self-driving labs, which are tools that can be used to generate high-quality experimental and simulation data.

In addition to these esteemed institutions, collaboration with **Artificial, Inc.** can greatly benefit the DOE by:

- **Streamlining Data Integration**: Our platform ensures seamless connectivity with diverse instruments and databases, facilitating a unified data environment that enhances AI readiness.
- **Advancing Automation and Efficiency**: Artificial automates routine procedures and workflows, reducing errors and freeing up researchers to focus on critical analysis and interpretation.
- **Providing Real-Time Insight**: By implementing real-time data monitoring tools, Artificial supports dynamic and adaptive experimentation, allowing for timely adjustments and insights into ongoing research processes.
- **Offering Customizable Solutions**: Our vendor-agnostic approach is adaptable to any laboratory setup, ensuring that DOE can leverage cutting-edge technology without significant infrastructure changes.

Partnering with **Artificial, Inc.** supports DOE’s strategic objectives to utilize AI-driven solutions for groundbreaking scientific research and energy innovations.

(d) What are additional data-related tools and technologies DOE should invest in to promote AI-ready data and fuel continued US leadership in AI?

To enhance DOE's AI research capabilities, Artificial, Inc. provides an integrated suite of tools designed to optimize data processing, analysis, and visualization within scientific laboratories. Key tools and their benefits include:

- **Digital Twin Platforms**: Utilizing Artificial's **Labs** application, DOE can create accurate digital replicas of physical labs. This facilitates real-time data integration, enabling seamless monitoring and control of laboratory operations to ensure data consistency and reliability.

- **Interactive Digital Assistants**: Artificial's **Assistants** transform standard operating procedures into interactive, digital workflows. This ensures standardized data collection and reduces human error, thereby enhancing data quality and integrity across experiments.
- **Workflow Orchestration Tools**: With **Workflows** and **LabOps**, Artificial integrates various lab instruments, software, and data sources into a unified platform. This promotes efficient data sharing and collaboration, allowing scientists to focus on analysis rather than data management.
- **Real-Time Data Scheduling and Monitoring**: **LabOps** provides a centralized interface for scheduling experiments and monitoring data streams in real-time. This ensures timely data acquisition and facilitates adaptive experimentation based on live data insights.
- **Extensible Adaptors for Data Integration**: Artificial's library of adaptors enables seamless connection to existing lab equipment and databases. This ensures comprehensive data integration, allowing the aggregation of diverse datasets for more robust AI model training and analysis.
- **Automated Data Logging and Provenance**: Every action within the lab is automatically logged and tracked, ensuring complete data provenance. This enhances data traceability and reproducibility, which are critical for validating AI-driven research outcomes.

By leveraging these advanced tools, DOE can ensure high-quality data management, foster collaborative research efforts, and accelerate the development of AI applications tailored to scientific needs. This alignment with Artificial, Inc.'s Product Suite supports DOE's objective of establishing a comprehensive national AI capability that drives innovation and serves the public interest.

Models

(a) To realize the benefits of open sourcing scientific and applied energy AI models, DOE must carefully balance the advantages of open collaboration with the requirement to protect national security and intellectual property.

Benefits of Open Sourcing:
Open sourcing AI models allows for a broader scientific engagement, accelerating innovation through cross-disciplinary collaboration and enabling more comprehensive peer reviews. The scientific community can rapidly validate and improve models, promoting transparency and advancing understanding across a wide array of energy-related challenges. For instance, in fields such as materials science and applied

physics, openness enables sharing of experimental data and algorithms, thus speeding up research cycles and fostering collaborations between national labs, industry, and academia. Such cross-sectoral collaboration aligns with the FASST initiative's goals of building the world's most powerful AI models for scientific and energy breakthroughs.

Research Security Considerations:
However, the release of advanced AI models and datasets, especially those containing sensitive or classified data, can introduce potential vulnerabilities. Striking the right balance between openness and security is essential. Models deployed for critical infrastructures like energy grids or used in national defense applications must undergo rigorous red-teaming and security audits to ensure robustness against adversarial exploitation. To that effect, DOE should adopt strict governance practices, such as ensuring that open-source models are sanitized of classified datasets, and implementing usage licenses that restrict exploitation by adversarial actors. DOE can also adopt a tiered-access model where specific production-ready AI models are made available to trusted partners only, while their derivative research or datasets can be more broadly shared with the academic and industrial communities.

Framework for Security:
DOE should establish standards for vetting contributors and parties accessing sensitive AI systems, involving measures such as background checks or whitelisting accredited organizations per defensive security protocols. Leveraging the governance practices developed by national labs and agencies like NIST in materials R&D, and integrating them with the latest AI safety protocols, could help create a secure, open-source ecosystem while retaining national interest.

ATI-like consortiums or controlled dissemination using platforms like GitHub with tiers of access, controlled by detailed license agreements or security certificates, could offer a structured avenue for DOE to manage this duality of openness and security.

(b) How can DOE support investment and innovation in energy efficient AI model architectures and deployment, including potentially through prize-based competitions?

To bolster energy-efficient AI models, particularly within **materials science** and **drug discovery**, DOE can adopt a strategic approach that merges public-private partnerships, cutting-edge research testbeds, and prize-based competitions. Here's how DOE can specifically nurture innovation:

1. Prioritize AI for **Materials Science** and **Drug Discovery**

- **Accelerate AI-Driven Research in Materials Science:** By focusing on self-driving laboratories and autonomous experimentation (AE), DOE can reduce processing times for new materials discovery, especially in fields like **high-entropy alloys** for aerospace or **advanced battery materials** for energy storage. These breakthroughs can be achieved by infusing AI models into the testbeds that simulate complex material behaviors.
- **Enable AI-Enhanced Drug Discovery Workflows:** AI models can drastically accelerate drug discovery by optimizing bioassay processes, chemical analysis, and synthesis. Through platforms like Artificial's **LabOps**, DOE can create highly efficient workflows that reduce time, cost, and resources needed to innovate.

2. Energy-Efficient AI Prize Competitions

- DOE can launch **prize-based competitions** to encourage industries, startups, and academic teams to develop energy-efficient AI architecture. These publicly funded challenges will incentivize the reduction of energy consumption in AI models that manage massive datasets. Example projects could include AI models powered by quantum computing or low-energy machine learning algorithms for **self-driving laboratories** used in materials and biotech R&D.
    - Competition Themes: "Scalable AI Models for Autonomous Materials Research" or "Energy-Efficient AI for Drug Optimization" could fund breakthroughs such as reducing neural network complexity linked to materials data analysis or drug compound screening.
    - Participation can include academic institutions (e.g., NIST, leading universities like University of Toronto), industry, and small to medium-sized enterprises, fostering interdisciplinary collaborations.

3. Autonomous Testbed Creation

- DOE should invest in "**Autonomous AI Testbeds**" within its national labs. These testbeds would serve to develop, validate, and optimize energy-efficient AI models in **materials science** and **pharmaceuticals**. Connected to industry and academia, these facilities could host prize-driven projects, allowing teams to iterate efficiently on physical prototypes with AI-based digital twins.

4. Standards Development & Modular Plug-and-Play Ecosystems

Leveraging systems like Artificial Inc.'s **Workflows** and **Labs**, DOE can work on establishing **AI standardization frameworks**, ensuring interoperability between diverse hardware, software, and AI models across industries. This will reduce redundant research and minimize energy-intensive custom solutions, fostering rapid scalability of energy-efficient AI applications in laboratories.

Thus, through these strategic methods, DOE will catalyze innovation in energy-efficient AI models, carving a path for sustainable AI advancements in these disruptive fields.

5. Moving Beyond R^2 Hacking

The community has reached a breakpoint whereby for a large number of AI for materials learning tasks, the existing models are sufficiently predictive inside of the distribution of the training set. The DOE should move towards investments that treat model optimization as a multi-objective optimization that uses a more nuanced view of model performance that accounts for interpretability, equity, training scores, validation scores, model stability, and robustness.

Applications

(a) What are application areas in science, applied energy, and national security that are primed for AI breakthroughs?

Artificial Intelligence (AI) breakthroughs are particularly primed in two crucial areas: **materials science** and **drug development**—both of which can be greatly enhanced through the application of self-driving laboratory methodologies.

1. Materials Science
AI and automation in materials research promise to shorten the discovery time for new materials, which is critical for advancing various industries such as energy storage, aerospace, and biotechnology. Self-driving labs equipped with autonomous experimentation (AE) platforms are particularly effective in this domain. Through the integration of AI models, these labs can rapidly explore massive parameter spaces, identify optimal material compositions, and refine manufacturing processes. AI can also accelerate the evolution of materials such as high-entropy alloys, composites, and photovoltaic materials by detecting hidden patterns in large experimental datasets. Self-driving labs that automate these processes could become foundational for managing the lifecycle of emerging materials, meeting the needs of critical applications in national security such as nuclear energy materials.

2. Drug Development
Self-driving labs bring immense potential in drug development, especially as they streamline the traditionally labor-intensive processes involved in bioassays, chemical synthesis, and compound screening. Integrating AI into these drug development workflows significantly optimizes **lead discovery**, **target identification**, and **toxicity screening** by interpreting vast datasets from numerous experimental trials. Artificial, Inc.'s **LabOps** platform enables real-time data integration and analysis, allowing researchers to make data-driven decisions rapidly. AI-powered models, used within

these self-driving labs, can also assist in **predicting molecular interactions** and refining synthesis protocols, drastically reducing time to market for critical pharmaceuticals and biotech innovations. Furthermore, incorporating AI-driven exploration and optimization algorithms supports adaptive experimentation, fostering continuous improvements in drug efficacy and safety.

Applying AI within self-driving labs to materials science and drug development will unlock opportunities for **dynamic experimentation** and discovery, enabling innovations that are central to overcoming current technological limitations and enhancing national security. These breakthroughs align closely with DOE's objectives in optimizing energy and scientific domains for substantial scalability and efficiency improvements.

(b) How can DOE ensure foundation AI models are effectively developed to realize breakthrough applications, in partnership with industry, academia, and other agencies?

The Department of Energy (DOE) can effectively ensure the development of foundation AI models to realize breakthrough applications by fostering collaboration between industry, academia, and government agencies through the following strategies, particularly in the realm of self-driving labs:

1. Public-Private Partnerships for Scalable AI Models

DOE should leverage its unique position to foster mission-driven public-private partnerships involving national labs, AI companies, and academic institutions. Self-driving labs, spearheaded by platforms like Artificial, Inc.'s **LabOps**, offer an opportunity to create AI-driven workflows that extend human expertise through automation and continuous feedback loops. By coordinating partnerships around key energy and scientific challenges, DOE can drive the co-development of scalable AI models. These partnerships can facilitate technology transfer, enabling the best practices from industry to merge with cutting-edge research from academia.

Example: DOE could extend its public-private partnership model by leveraging platforms like Artificial, Inc.'s **LabOps** to integrate AI orchestration with automation across its labs and national facilities. Artificial's solutions, particularly in creating **Digital Twins** of labs and automating complex workflows, can align seamlessly with DOE's objectives, creating an AI-driven ecosystem that interlinks advanced laboratory systems via unified, cloud-based infrastructure.

2. Establishment of Autonomous Testbeds

DOE should accelerate the creation of autonomous laboratories and **self-driving testbeds** in national labs in collaboration with industry and academic partners. Establishing such testbeds allows for rigorous experimentation and validation of models in controlled but real-world scenarios. Such environments can fuel AI breakthroughs in areas like materials science, energy storage, and pharmaceuticals, by allowing continuous AI-driven experimentation.

Example: By incorporating platforms like Artificial, Inc.'s **Digital Twin Labs**, these testbeds could operate with minimal downtime, generating high-quality, AI-ready datasets that validate foundation AI models in energy-related tasks.

3. Standardization and Interoperability Framework

Self-driving labs require the robust integration of hardware, software, people, and data streams. DOE should work with partners such as NIST to establish standards for data sharing, instrument control, and AI integration within autonomous labs. Standardization lowers barriers to AI and automation adoption, ensuring that industry partners can more rapidly implement self-driving lab technologies, while ensuring data interoperability and security.

Example: The **Workflows** and **Assistants** platform from Artificial, Inc., which already standardizes lab protocols, could be extended to national labs to synchronize AI workflows across diverse lab environments.

4. Competitive Funding and Directed Incentives

DOE should promote prize-based competitions and directed funding initiatives to expedite the development of energy-efficient AI models and frameworks critical to self-driving labs. By incentivizing ecosystem development, including the creation of plug-and-play components for autonomous systems, DOE can fill critical gaps in the current AI infrastructure.

Example: Directed funding, like SBIR grants, could support startups and academia in AI tool development. These grants would enhance platforms like **Artificial**, improving AI integration in self-driving labs and expanding lab capabilities to meet DOE's complex demands. This would make AI-driven automation more accessible, boosting economic growth and technological leadership in critical sectors.

By combining public-private collaboration, autonomous testbed development, standardization, and funding mechanisms, DOE will lay the foundation for breakthrough AI model development within self-driving labs to propel the U.S. into a leading position in AI-driven research across energy, science, and national security.

https://www.federalregister.gov/documents/2024/09/12/2024-20676/notice-of-request-for-information-rfi-on-frontiers-in-ai-for-science-security-and-technology-fasst