Testimony of Dr. Georgia Tourassi

Associate Laboratory Director for Computing and Computational Sciences

Oak Ridge National Laboratory

Before the

Joint Research & Technology and Energy Subcommittees

Science, Space, and Technology Committee

U.S. House of Representatives

February 6, 2024

Hearing on "Federal Science Agencies and the Promise of AI in Driving Scientific Discoveries"

Chairman Collins, Chairman Williams, Ranking Member Stevens, Ranking Member Bowman, and distinguished members of the Committee: Thank you for the opportunity to appear before you today. My name is Georgia Tourassi. I am the Associate Laboratory Director for Computing and Computational Sciences at the U.S. Department of Energy's Oak Ridge National Laboratory (ORNL) in Oak Ridge, Tennessee. In the past four years, I have been privileged to direct one of the Department of Energy Office of Science User Facilities, the Oak Ridge Leadership Computing Facility (OLCF), and help deploy Frontier, the world's most powerful supercomputer dedicated to open science. It is an honor to provide this testimony on the transformative role of the Department of Energy's efforts in spearheading Artificial Intelligence (AI) advances to support its mission in science, energy, and security, giving the nation a significant advantage in the ever-changing and intensifying scientific and economic landscape.

I am a biomedical engineer and computational scientist by education and training. My research is at the intersection of AI, high-performance computing (HPC), and biomedicine. The views I offer in this statement are my own, shaped by my scientific experiences as an AI translational scientist for the past three decades, my experience with inter-agency partnerships to advance respective missions, and my experience as Director of a Department of Energy (DOE) Leadership Computing Facility dedicated to open science.

I will begin with a brief overview of the Department of Energy's leadership in high performance computing and how it establishes the essential foundation for U.S. competitive leadership in AI. Then, I will describe DOE's efforts in developing and responsibly applying AI solutions to serve the department's mission in energy, science, and security while understanding and balancing the risks and benefits of accelerating AI. Lastly, I will discuss how DOE supports the National Artificial Intelligence Research Resource (NAIRR) goal and the importance of a coordinated U.S. strategy in AI.

DOE Spearheading Leadership Computing to Ensure U.S. Global Primacy in AI

DOE has historically played a major role in spearheading critical infrastructure and technologies for accelerating the pace of scientific innovation in the United States. Through its network of national labs and world-renowned scientific user facilities, the DOE has been leading foundational research and applied technology development, keeping the United States competitive across scientific frontiers most critical for the 21st century's economy and security.

Specifically, high-performance computing (or supercomputing) has been the cornerstone for research advancement, scientific innovation, competitive advantage, and economic prosperity[1]. The Department of Energy's Leadership Computing Program has a long history of enabling researchers to accelerate scientific discovery and deliver practical breakthroughs for some of the most computationally challenging problems.

For example, the Oak Ridge Leadership Computing Facility (OLCF) has been at the forefront of the rapid evolution in high performance computing for decades. The OLCF provided the first teraflop system (IBM Power3 Eagle) for open science and the

[1] US DOE Office of Science. High-performance Computing. https://www.energy.gov/science/initiatives/high-performance-computing. Accessed January 31, 2024

science community's first petaflop system (Cray XT5 Jaguar), both of which are considered major historical milestones in computing. The last four consecutive OLCF systems ranked as the fastest in the world, including the first exascale system in the U.S., the 1.2 exaflops Frontier supercomputer which debuted in 2022.[2] **Frontier marked an impressive trajectory for our facility, for DOE, and for the nation,** demonstrating that in the past decade we improved computational power by a factor of 500 while suppressing growth of energy use, or the carbon footprint, which increased only by a factor of 4. We achieved this in partnership with U.S. vendors and have been at the forefront of adopting and enhancing GPUs to reduce energy consumption for science and engineering applications.

Sustaining U.S. leadership in high performance computing is critical for preserving American competitiveness in AI. The immense computing power required to train and optimize complex AI models such as generative AI large-language models increasingly demands computing architectures that can deliver both cutting-edge performance and optimal energy efficiency. As AI models grow ever larger and more computationally intensive, the energy costs and demands will balloon without continued investment in efficiency gains. With international competition from China and other nations intensifying, advancing post-exascale supercomputing can give the nation a competitive advantage. Investing in hardware innovations around novel chiplets, optimized software, and integrative computing paradigms can drive efficiency gains and provide U.S. researchers and industries with a competitive edge to continuously advance AI. DOE has been making vital R&D investments strategically targeting these areas to extend the transformative gains achieved through the Exascale Computing Project (ECP).[3]

DOE Leadership Computing Enables Large-Scale AI Applications

Today, all science is computational science, and increasingly, more scientific domains explore and leverage AI. DOE's leadership computing facilities are proving essential to train, iteratively refine, and productively deploy the complex algorithms and

[2] https://www.ornl.gov/news/frontier-supercomputer-debuts-worlds-fastest-breaking-exascale-barrier, Accessed January 31, 2024
[3] https://www.exascaleproject.org/, Accessed January 31, 2024

the largest scale models underlying modern AI breakthroughs in academic and government research and industry partners.

Frontier supports researchers from academia, industry, and federal agencies in **training AI models with millions and often billions of parameters** to advance science and offer new insights into quantum materials, chemistry, physics, biology, additive manufacturing, fuels, fusion, fission, and more. Here are some examples of how AI scientific campaigns on Frontier are already making an impact.

Climate modeling: A team of computational scientists at ORNL is training trustworthy AI foundational models using a large volume of simulation and observational climate data on Frontier. These models are specifically designed to effectively capture both spatial and temporal patterns inherent in large-scale climate and weather datasets. These models will provide rapid and accurate predictions for a broad range of climate and weather forecasting tasks and inform decision-making.

Optimizing the efficiency of fusion reactors: Scientists from the Princeton Plasma Physics Laboratory and Brookhaven National Laboratory are using Frontier to train foundational AI models on data sourced from large-scale simulations and actual data from fusion reactors. These models will be used to predict disruptions within the reactors, thereby enabling proactive plasma control. Such insights are crucial to enhance the efficiency of fusion reactors.

Accelerating drug discovery: A team of computational scientists at the Oak Ridge National Laboratory used Frontier to design a machine-learning software stack that predicts how strongly a given drug molecule will bind to a pathogen as well as how it will attach to the target. These vital pieces of information can greatly shorten the usual trial-and-error process of lab experimentation to find viable drug candidates, especially for novel viruses like COVID-19 with yet unknown structures.[4]

Deploying AI solutions to fight cancer: Supporting efforts to reduce the cancer burden in the United States, the National Surveillance, Epidemiology, and End Results (SEER) Program manages the collection of curated data from population-based cancer

---

[4] https://www.olcf.ornl.gov/2022/11/11/predicting-protein-targets-and-the-molecules-that-bind-with-them/, Accessed January 31, 2024

registries on cancer incidence, prevalence, survival, and associated health statistics for the advancement of public health. As part of the National Cancer Institute (NCI), the SEER program supports cancer research to improve the understanding of patient care and outcomes in the “real world” beyond the clinical trial setting. By leveraging ORNL’s Frontier system, a DOE-NCI research team delivered state-of-the-art AI tools to help US state cancer registries improve the accuracy and efficiency of their operations in cancer incidence reporting. The AI tools make the dream of “real time” cancer incidence reporting a reality, while also enabling real-time eligibility assessment of cancer patients for clinical trials. This achievement to modernize the national cancer surveillance program exemplifies the benefits of a federally coordinated strategy to harness AI, high-performance computing, and sensitive health data assets for real-world application.[5]

Fast-tracking medical discovery: Using Frontier, a multi-institutional team of academic, DOE, and industry researchers scanned hundreds of thousands of biomedical concepts from millions of scientific publications in search of potential connections among symptoms, diseases, conditions, and treatments. The team developed an AI-based method capable of pinpointing potential links amid millions of medical concepts across decades of scientific publications. Their work on Frontier reached a speed of 1 exaflops at single precision and took only 11.7 minutes to search more than 7 million data points drawn from 18 million publications. The effort began as part of the fight against COVID-19 but could eventually become as essential to basic diagnosis and treatment as Google is to an online search.[6]

Accelerating materials research: ORNL researchers developed AtomAI, an AI tool to provide end-to-end image analysis of electron and scanning probe microscopy images. AtomAI applies deep learning to microscopy data at atomic resolutions, thereby providing quantifiable physical information such as the precise position and type of each atom in a sample. AtomAI enables researchers to quickly derive statistically meaningful information from immensely complex datasets that routinely include hundreds of images, each containing thousands of atoms and abnormalities in molecular structure.

[5] https://www.ornl.gov/news/researchers-honored-innovative-use-ai-fight-cancer, Accessed January 31, 2024
[6] https://www.olcf.ornl.gov/2022/10/25/fast-tracking-medical-discovery/ Accessed January 31, 2024

AtomAI allows researchers to engineer atomically precise abnormalities in materials and gain deeper insights into the materials' physical and chemical qualities.[7]

Energy-efficient training of large language models: A team led by ORNL researchers used Frontier to investigate training strategies for one of the largest AI models to date. The team ran the initial stages of training on a large language model similar to OpenAI's ChatGPT or Together's RedPajama 7B. The study ran models with up to 1 trillion parameters on 1,024 and 3,072 of Frontier's more than 9,400 nodes focusing **on pinpointing the most efficient ways to exploit Frontier's GPUs** for AI training. The researchers identified that a blend of data parallelism and porting training strategies performed best when tailored to the computing platform.[8]

DOE Advancing Mission Critical AI

DOE is keenly aware of the importance of AI in advancing its mission in science, energy, and security. DOE and its national laboratories have invested in AI development and applications since the early 1960s, delivering cutting-edge AI tools and data science capabilities to complement and leverage its leadership in high-performance computing.

In 2023, DOE issued the report "AI for Science, Energy, and Security"[9] which was the culmination of a series of "townhalls" engaging over 1,300 scientists and engineers from all 17 national laboratories, more than 30 universities, and dozens of companies to identify guiding principles and strategies for advancing AI to achieve DOE's mission objectives. The report identified several **mission-critical AI applications of national importance**:

- Science: Use of AI to accelerate scientific discovery and innovation in areas like materials, climate, biology, and nuclear science. Important applications include AI to analyze the vast volumes of experimental data generated in DOE's user

[7] https://www.ornl.gov/news/deep-learning-based-data-analysis-software-ornl-promises-accelerate-materials-research, Accessed January 31, 2024
[8] https://arxiv.org/abs/2312.12705, Accessed January 31, 2024
[9] https://www.ornl.gov/file/advanced-research-directions-ai-science-energy-and-security/display, Accessed February 1, 2024

facilities, to model complex phenomena like climate and biological systems, to design new materials and chemicals, and to optimize research workflows.

- Energy: Use of AI for more efficient, resilient, and secure energy systems including the grid infrastructure. Important applications include AI to optimize complex energy systems, to simulate and control fusion reactions, to manage smart grids, to model fuel sources, to enable autonomous platforms for monitoring, and more.
- Security: Use of AI to enhance national security by pioneering important AI solutions in critical areas such as nuclear nonproliferation, cybersecurity, intelligence analysis, emergency response, and resilience against natural and man-made threats.

To deliver on the strategic vision outlined in the report, the DOE national labs collaborated to develop a programmatic roadmap to deliver AI-enabled solutions and mission-critical AI applications in energy, science, and security. The Frontiers in Artificial Intelligence for Science, Security, and Technology (FASST) roadmap is structured around multi-lab AI research centers with thrust areas that are foundational to advance AI applications in energy, science, and security such as AI foundation models for scientific discovery, AI prediction and control of complex engineered systems, AI for software engineering and computer programming, and others described in the 2023 DOE report.

The FASST research centers are envisioned as engines driving AI progress and amplifying innovation through open partnerships with academia and industry, leveraging DOE's unique compute resources, large scientific data assets, multi-disciplinary domain expertise, and **deep experience in translating effective solutions from open science to national security.**

DOE Advancing Energy-Efficient, Secure, and Trustworthy AI

Aside from mission-relevant applications, the DOE report identified **critical research** needed to improve AI performance, energy consumption, and methods for ensuring safety, reliability, and trustworthiness. When democratizing access to powerful but still maturing technologies like AI, careful consideration must be given to balancing risk mitigation with accelerating innovation. Research focused on performance, safety, efficiency, and trust can inform policies and investments that optimize public benefit.

Proliferating AI capabilities are fueling creative new applications and breakthroughs at a pace that is not feasible when that expertise and tools are concentrated in the hands of a few. But rapid, uncontrolled democratization also courts perils like inadequate testing, misuse, embedded biases, and risky overreliance on still unproven and opaque tools. Strategic R&D investments and systematic approaches are thus crucial to realize democratization's upside while de-risking downsides.

Keeping pace with rapidly changing AI breakthroughs requires a dedicated focus on "regulatory science" – research focused on the methods and policies to assess the safety, effectiveness, trustworthiness, and other oversight questions raised by new technologies. Risk analytics, uncertainty quantification, surveillance capabilities, predictive computational models, and integrative assessment methods underpin evidence-based regulation. DOE's Advanced Scientific Computing Research (ASCR) program has long supported applied mathematics research in these areas. Such research is needed to counteract the risks of irresponsible AI applications or adversarial use. Leadership computing facilities supported by ASCR have a long history of leveraging multidisciplinary teams and applying safety practices founded on technical, administrative, and policy solutions that are continuously evaluated and refined to meet emerging operational challenges. Support for open benchmarks, shared standards, and access to reference data grounds DOE's validation, verification, and deployment of technical solutions in real world settings. The FASST vision builds on DOE's experience with a sharpened focus on the crosscutting themes of AI performance, energy **efficiency, safety, and trustworthiness**.

DOE Supporting the NAIRR Pilot Program

Throughout the years, DOE lent its expertise in major scientific user facilities and advanced research capabilities to **multi-agency collaborative projects** advancing shared national priorities around health (e.g., Human Genome Project, Cancer

Moonshot), infrastructure (e.g., Grid Modernization), technology strength (e.g., National nanotechnology initiative), and crisis response (COVID-19 Response) to name a few. Currently, DOE's Office of Science is contributing **significant resources** and data assets to the National Artificial Intelligence Research Resource (NAIRR) pilot, a concept for a shared national research infrastructure to connect U.S. researchers to responsible and trustworthy AI resources.

Specifically, **DOE contributes to the NAIRR pilot by providing access to the extreme scale computing capabilities of the Oak Ridge Leadership Computing Facility's Summit supercomputer.** Currently ranked the seventh fastest supercomputer in the world, Summit is a powerful and reliable instrument for AI-driven scientific discoveries in energy, climate, health, and other areas. In partnership with IBM and NVIDIA, DOE extended Summit operations for a year to support AI-enabled open science at a supercomputing scale. Because Summit also has Authority to Operate (ATO) with sensitive data, it is an environment ideally suited for the **"NAIRR Secure"** focus area of the NAIRR pilot, co-led by NIH and DOE, to enable AI research with personal and proprietary data requiring privacy protections and security.

As part of its contribution to the NAIRR pilot, DOE also provides access to the Argonne Leadership Computing Facility's AI Testbed. The Argonne Leadership Computing Facility (ALCF) AI Testbed is a growing collection of some of the world's most advanced AI accelerators available for open scientific research. The ALCF offers access to its Cerebras, Graphcore, Groq, and SambaNova systems and training to leverage these resources effectively. In addition, as a leading producer of scientific data, DOE is contributing unique data assets from materials science, ground-based atmospheric measurements for advancing climate science, nuclear data, and genomic data for advancing systems biology.[10]

[10] https://nairratdoe.ornl.gov/, Accessed January 31, 2024

AI Leadership Demands Investments that Balance Risk and Opportunity, Basic and Applied Science, and Coordinated Efforts Among Federal Agencies

Interest in AI appears to be dominated by practical applications with direct societal impact, particularly with respect to protecting our nation's security. Such a focus is a reasonable strategy when there are multiple competing priorities and finite resources, but it can also be our Achilles' heel. Many impactful applications of AI build on years or decades of more fundamental, curiosity-driven research. An overemphasis on immediately usable AI applications risks overlooking investments needed now to understand potential long-term or unintended consequences of AI. Curiosity-driven research can lead to major unexpected breakthroughs and new insights that form the basis of transformative applications. These can't be planned for in advance.

A seminal example of fundamental science to major breakthrough in DOE's history comes from basic nuclear physics research enabling a myriad of later technological applications empowered by the discovery of nuclear fission in the 1930s and 1940s. At the time, the splitting of uranium atoms into lighter nuclei was a purely scientific breakthrough, revealing new subatomic properties and reactions. The researchers undertaking early fission experiments were motivated by curiosity about fundamentals of the nucleus rather than any practical uses.

Yet this foundational breakthrough laid the groundwork for the later development of nuclear reactors for energy production, medical isotopes for imaging and cancer treatment, radiation therapies for attacking tumors, dating techniques for archaeological artifacts, radioisotope thermoelectric generators that power space exploration, and even nuclear weapons as a deterrent to global war. None of these modern technologies would have been possible without the original basic science research into nuclear fission done out of academic curiosity rather than application-focused goals. And many of the resulting applications could not have been envisioned when fission was first detected in the laboratory.

This case encapsulates how seemingly obscure basic science discoveries branch out over time into diverse practical uses far beyond their initial scope. Investing in AI to accelerate basic scientific discovery even when applications are unclear is essential, as one cannot predict where fundamental breakthroughs may ultimately lead further down the road. Federal R&D investments in AI need to nurture basic and applied science.

While both NAIRR and FASST aim to push forward AI capabilities to accelerate science, the scope and intricacy of the challenges they plan to tackle have key distinctions. NAIRR takes a broad approach, providing infrastructure to empower the AI research community to make foundational advances across a wide range of scientific domains. NAIRR focuses on building **AI capacity** for the nation. In contrast, FASST aims on building **AI capability** for the nation by focusing intensely on complex mission-driven challenges in DOE-relevant areas which require the coordination of integrated teams leveraging exascale-class computing resources. While the NAIRR challenges are critically important for expanding the frontiers of AI, the FASST challenges exist at a greater level of complexity and scale, and with a high bar for managing risk due to DOE's unique mission focus. Together these complementary but different initiatives work to comprehensively advance AI for the benefit of science and the nation.

CONCLUDING REMARKS

As a science and technology powerhouse, the DOE national laboratories possess **unparalleled compute and data resources** and **multi-disciplinary talent pools** to advance cutting-edge artificial intelligence **from research to fielded applications in challenging mission domains**. With their strong history in fostering public-private innovation, close collaborations with academia, and strategic partnerships with other federal agencies, DOE national laboratories drive scientific research and the transfer of the resulting value to the nation and society.

Broadly empowering society via technology requires enlightened restraint as much as unrestrained ambition. To ensure AI serves our best interests as its influence grows, improving its performance and energy-efficiency must be responsibly balanced, maximizing its safety, reliability, and transparency. DOE has demonstrated how to **balance risk and opportunity** to deliver solutions with broad and lasting impacts on science and the nation. **Leveraging DOE's capabilities will be key to the United States maintaining its edge in the AI race unfolding worldwide.**

Thank you again for the opportunity to testify. I welcome your questions on this important topic.