The Honorable Michael J.K. Kratsios

Fourth Chief Technology Officer of The United States
The White House

Former Under Secretary of Defense for Research and Engineering (A)
United States Department Of Defense

Managing Director
Scale AI

Distinguished Fellow
Council on Competitiveness

Balancing Knowledge and Governance: Foundations for Effective Risk Management of Artificial Intelligence

Statement before the Science, Space, and Technology
Joint Investigations & Oversight and Research & Technology Subcommittees
United States House of Representatives

Washington, D.C.

October 18, 2023

Chairman Obernolte, Chairman Collins, Ranking Member Foushee, Ranking Member Stevens, and Members of the Investigations and Oversight and Research and Technology Subcommittees, thank you for the opportunity to be here today to testify on the critical topic of developing safe, secure, and responsible artificial intelligence.

I am honored to be here today to discuss these topics with you.

My name is Michael Kratsios, and I am the Managing Director of Scale AI, an artificial intelligence company headquartered in San Francisco. Prior to my time at Scale, I served as the fourth Chief Technology Officer of the United States at the White House and also as the Acting Under Secretary of Defense for Research and Engineering at the Pentagon.

While at the White House, I helped architect the country's first national AI strategy, the American AI Initiative.

Scale, where I work today, was founded in 2016 with the mission of accelerating the development of AI. Scale works with leading frontier model builders like OpenAI and Meta, as well as with the federal government, including the Department of Defense's Chief Digital and AI Office and the U.S. Army.

Today, Scale fine-tunes, red teams, or tests and evaluates nearly all of the leading frontier large language models (LLMs). This gives Scale a unique vantage point to understand and advance the development of safe, secure, and trustworthy AI.

* * *

I'd like to use my statement to make two points.

First, the advent of large language models represents a turning point in artificial intelligence. However, as remarkable as these systems are, they are not our first encounter with the promise and risks of AI. For years, policymakers have grappled with AI oversight. The prudent path forward relies on building on existing efforts and adopting use case and sector-specific, risk-based guidelines for responsible AI deployment.

Since 2016, successive administrations and Congress have laid a robust foundation for AI policy and regulation. This includes a report on the future of AI under President Obama and the nation's first AI strategy, the American AI Initiative, and multiple executive orders under President Trump.

The Trump Administration made AI a national technology priority. It committed to doubling federal AI research spending,[1] created first-of-a-kind national AI research institutes,[2] developed the world's first guidance for the regulation of AI in the private sector,[3] and collaborated with allies to develop the world's first intergovernmental AI policy guidelines at the OECD.[4]

Congress codified many of these efforts in the 2020 National AI Initiative and AI in Government Acts, including the establishment of the National AI Initiative Office at the White House to coordinate AI policy efforts across agencies and offering a blueprint for the creation of a National AI Research Resource (NAIRR).

The Biden Administration has continued work on AI, releasing a Blueprint for an AI Bill of Rights[5] and earlier this year securing voluntary commitments from a number of technology companies regarding AI safety.[6] In January of this year, NIST released the government's first AI Risk Management Framework, in fulfillment of another directive of the National AI Initiative Act of 2020.[7]

Unfortunately, implementation of existing legislation and the 2019 and 2020 Executive Orders has fallen short. A study by Stanford University's Institute for Human Centered AI found that less than 40 percent of the 45 legal requirements associated with the AI in Government Act and the two AI Executive Orders have been implemented.[8] Notably the Office of Management and Budget has yet to publish required guidance to support agency use and acquisition of AI.

While considering additional legislation or pursuing new administrative actions on AI, policymakers should first ensure that federal agencies fully implement existing laws and are following through with requirements from executive orders relating to responsible AI adoption.

When it comes to regulating AI systems, lawmakers should pursue a use case and sector-specific, risk-based approach rooted in high quality testing and evaluation. This approach tailors the rigor of evaluation to the level of risk posed by each use case. For example, AI-powered medical diagnostics, which carry higher risks to human health and safety, should undergo more stringent testing and evaluation by the FDA compared to lower-risk AI applications like a movie recommendation algorithm.

---

[1] https://www.wsj.com/articles/trump-wants-to-double-spending-on-ai-quantum-computing-11581378069
[2] https://www.wsj.com/articles/white-house-announces-1-billion-plan-to-create-ai-quantum-institutes-11598432400
[3] https://trumpwhitehouse.archives.gov/wp-content/uploads/2020/11/M-21-06.pdf
[4] https://legalinstruM-21-06.pdfments.oecd.org/en/instruments/OECD-LEGAL-0449
[5] https://www.whitehouse.gov/ostp/ai-bill-of-rights/
[6] https://www.whitehouse.gov/briefing-room/statements-releases/2023/09/12/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-eight-additional-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/
[7] https://www.nist.gov/itl/ai-risk-management-framework
[8] https://dho.stanford.edu/wp-content/uploads/AI_Implementation.pdf

We are not starting from scratch on this approach. Test and evaluation is a standard part of bringing products to market in nearly every industry. With regards to AI risk and regulation, current OMB guidance to agencies on regulating AI takes a use case specific approach and is very clear. "A risk-based approach should be used to determine which risks are acceptable and which risks present the possibility of unacceptable harm, or harm that has expected costs greater than expected benefits."[9] I believe this guidance remains fair and appropriate.

The Biden Administration has recognized the value of red teaming and test and evaluation, both in their voluntary commitments and through their support for the DEF CON31 AI Village Red Team event in August of this year. The event brought together over two thousand participants to red team eight leading LLMs on a test and evaluation platform built by Scale. This event demonstrated the critical role that test and evaluation plays in both model development and ensuring AI is safe to deploy.

In short, there has been a tremendous amount of work on AI oversight to date across the Federal Government. Any new legislative and regulatory actions should build upon and supplement this existing robust body of work and should focus on ensuring a use case and sector-specific, risk-based approach to regulation.

* * *

Second, to ensure the development and deployment of safe, secure, and responsible AI, we must harness the full strength of America's innovation ecosystem. America's unique innovation ecosystem of government, industry, and academia has collectively driven countless technological advances that have shaped the modern world.

In the past, the federal government has focused resources on funding early stage, basic, pre-competitive research and development, while visionary industry leaders commercialized those discoveries for broad public benefit. However, rapid advancements in state of the art AI are being propelled today primarily by private companies. This landscape requires increased ecosystem collaboration to ensure we develop AI that is safe, secure, and responsible.

Industry is participating in a number of consensus-based forums, like the Frontier Model Forum and ML Commons, to address gaps in and conduct research and develop on AI standards, such as those for testing and evaluation, which still are a technical work in progress.

Scale recently published its own technical methodology for LLM test and evaluation, which combines automated assessment with human evaluation of AI systems.[10] However, there is still

[9] https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf
[10] https://scale.com/guides/test-and-evaluation-vision;
https://static.scale.com/uploads/6019a18f03a4ae003acb1113/test-and-evaluation.pdf

much work to be done, and collaboration among industry, the government, and academia is essential.

The federal government must continue driving this innovation. In particular, the Department of Energy's national laboratories represent unique strategic assets, both for fundamental AI research, but also very importantly for assessing potential risks. With their world-leading supercomputing capabilities, multidisciplinary expertise, and university partnerships, DOE could be a natural home for the U.S. Government's work on assessing frontier risk.

Additionally, Congress has the opportunity to take a major step forward by formally establishing and funding the NAIRR. Doing so would create a critical shared research infrastructure, providing expanded access to computational capabilities, high-quality datasets, and educational resources. Access to high-quality training data is the bedrock for developing responsible AI systems. The NAIRR could serve as a crucial engine to drive the curation and availability of such data across government agencies.

Beyond research, the government should convene stakeholders, provide direction, and assert American leadership at international standards organizations. The United States should continue to engage with our partners and allies across relevant fora like the G7, G20, and the OECD to ensure our American ideals and values continue to underpin the development of AI. We must not cede this ground to authoritarian governments who do not share our same values.

Our nation's universities must continue tackling complex research problems and lend their expertise to collaborative working groups. Robust, federally funded research will remain critical to advancing the state of the art in a responsible manner.

Only by marshaling the full strength of the American innovation ecosystem can we ensure American leadership in developing trustworthy AI.

* * *

I truly believe that the United States can maintain and expand our global leadership in artificial intelligence. American leadership in AI will be critical for maintaining the economic and national security of the United States.

Thank you again for the opportunity to be here today, and I look forward to your questions.