**Frontier Model Forum Response to the Request for Information on the Development of an Artificial Intelligence Action Plan**

*The Frontier Model Forum submitted the below response on March 14, 2025 to the Request for Information on the Development of an Artificial Intelligence (AI) Action Plan.*

We are grateful for the opportunity to respond to the request for information from the Office of Science and Technology Policy (OSTP) on the "Development of an AI Action Plan."

The Frontier Model Forum is an industry-supported[1] non-profit dedicated to advancing the secure development and deployment of frontier AI systems. We leverage the technical and operational expertise of our member firms, as well as the broader scientific community, to develop industry best practices for managing the most significant national security and public safety risks related to frontier AI systems.

Our efforts focus primarily on chemical, biological, radiological, and nuclear (CBRN) and advanced cyber threats. We have dedicated workstreams for securing frontier AI models and systems, for addressing novel cyber capabilities that could amplify existing threats or create new ones, and for advancing the biosafety and biosecurity of frontier AI. Each of our workstreams aims to develop shared understandings of relevant threat models, safety evaluations, and mitigation measures. We take a rigorous, evidence-based approach that is consistent with scientific findings to identifying and mitigating credible risks.

We firmly believe that AI has the potential to benefit society in transformative ways, but that to harness its benefits we must successfully manage potential national security and public safety risks. Below we offer several priority issue areas related to our work that may advance the AI Action Plan developed by OSTP.

**Supporting the science of AI security and safety**

The science of AI security and safety is still nascent. Many open questions remain about how best to test and evaluate the advanced capabilities of frontier AI models and systems, as well as how to identify and mitigate their potential risks.

To address those gaps, the U.S. should continue to support and invest in:

[1] Our current member firms are Amazon, Anthropic, Google, Meta, Microsoft, and OpenAI.

- **AI metrology and measurement**. Developing rigorous, replicable and tailored research designs for evaluating advanced AI systems is essential for understanding their potential benefits and risks with respect to national security and public safety.

- **AI risk management**. Developing robust processes for managing the national security and public safety risks of frontier AI systems - including procedures for modeling, evaluating, and mitigating high-risk threats - is also vital.

- **Domain expertise**. Advancing the security of frontier AI systems requires deep expertise in AI as well as relevant risk domains. The administration should work to ensure that efforts to improve AI security and risk management include relevant subject matter experts with relevant scientific expertise.

The National Institute of Standards and Technology (NIST) and other research agencies, such as the National Labs in the Department of Energy, are well-positioned to help refine the emerging measurement science of evaluations. NIST is also well-equipped to develop robust processes for responsibly managing the national security and public safety risks of frontier AI systems.

**Strengthening international coordination and global standards**

As the first Trump administration recognized, the U.S. has a vital role to play in developing and championing scientifically informed standards for AI that balance innovation with appropriate safeguards for high-risk applications.[2] Strong U.S. leadership in setting international AI standards will reduce fragmentation and foster global alignment on security and governance practices that reflect democratic values.

The U.S. government should also strengthen coordination with foreign government agencies and offices tasked with evaluating advanced AI systems, such as the U.K. AI Security Institute (AISI) and related bodies. Engaging with foreign AISIs will be essential for ensuring effective and secure global deployment of frontier AI systems developed in the U.S. and for strengthening international trust. From a research perspective, it can contribute to identifying evidence gaps, developing more effective risk mitigations, and synthesizing research and applied learnings.

**Advancing national security testing and coordination**

The U.S. government possesses unique expertise and intelligence on CBRN and related national security issues that does not exist outside of government. At a moment when

[2] February 11, 2019, Executive Order on Maintaining American Leadership in Artificial Intelligence.

the science underpinning national security evaluations of frontier models is evolving quickly, industry efforts to evaluate models for national security risks benefit from close public-private collaboration, including programs that facilitate voluntary pre- and post-deployment evaluations by government experts.

The U.S. government should continue to support national security testing and evaluation, including the creation of testbeds at relevant bodies such as Sandia National Laboratories. In line with recent AI National Security Memorandums, it should also continue to coordinate across relevant national security agencies to facilitate information-sharing and develop integrated approaches to managing the national security risks related to advanced AI systems. Greater testing and evaluation capacity would also enable increased strategic awareness of how foreign models could be leveraged.

We appreciate the opportunity to comment and hope our recommendations prove useful. We look forward to supporting continued efforts by the U.S. government to advance secure, safe, and innovative AI development.

*This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in developing the AI Action Plan and associated documents without attribution.*