CHAMBER
OF PROGRESS

September 9, 2024

Department of Commerce
National Institute of Standards and Technology
U.S. AI Safety Institute
100 Bureau Drive
Gaithersburg, Maryland 20899

Re: NIST AI-800-1, Managing Misuse Risk for Dual-Use Foundation Models

On behalf of Chamber of Progress—a tech industry association supporting public policies to build a more inclusive country in which all people benefit from technological advances—we appreciate the opportunity to share this response to the U.S. AI Safety Institute at the National Institute of Standards and Technology's (NIST) Request for Information on Managing Misuse Risk for Dual-Use Foundation Models.

The United States is the global leader in AI. As NIST considers its approach to AI, it should prioritize preserving America as the epicenter of AI innovation. Overly prescriptive frameworks stifle innovation by creating high barriers for startups and smaller firms, limiting their ability to compete and innovate. NIST should support open access, encourage innovation, and avoid excessive constraints, thereby supporting a vibrant, competitive AI ecosystem.

We further caution that many of the purported risks of AI are not related to model development. Rather, the most cognizable - indeed extant - risks of AI are misuse of AI tools by third parties, including deceptions via deepfake audio or video.

An abundance of foundation models promotes competition
No evidence exists that open weight models are uniquely risky. Accordingly, we urge you to rethink Practices 3.1 and 3.3. Specifically, Recommendation 4 of Practice 3.3 encourages developers to "**Apply appropriate protections against insider threats, such as limiting access to model weights within the organization.**"[1] This recommendation is not founded; NIST has not established that access to model weights is risky, much less what mitigations might be necessary.

An abundance of foundation models, both open and closed, is vital for maintaining competition in the technology sector. Policy should encourage the proliferation of models

[1] NIST p.8, *Managing Misuse Risk for Dual-Use Foundation Models*, Dept. of Commerce (July 2024). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-1.ipd.pdf

- since models have different use cases depending on their size, design, etc. Additionally, fostering open source models will provide a base for researchers and developers to build upon.[2] Industry is leading the way. For example, Anthropic already offers an application programming interface (API) called Claude that provides access to advanced AI models designed for researchers to integrate into their projects and explore new applications.[3]

Foundation models benefit innovation
Foundation models, including open source models, drive technological innovation and competition by democratizing access to cutting-edge AI technology. By making the model architecture, code, and often the training data publicly available, these models enable a diverse range of developers and researchers to experiment and innovate. For example, Bidirectional Encoder Representations from Transformers or BERT is an open source language natural processing model developed by Google and is beneficial to developers and researchers because it provides a pre-trained model that can be fine-tuned for a range of tasks such as question-answering and language translation. By building on BERT, developers can save time and resources, and researchers can use it as a foundation for exploring new methodologies and accelerating progress in the field.[4]

However, such benefits will be harmed under Practice 4.1, specifically Recommendation 3, which encourages developers to "**Consider measuring model performance on proxy tasks that are safe and tractable while being similar enough to allow reliable inferences about the capability of concern.**"[5] This recommendation underscores the complexity of evaluating foundation models, suggesting that using proxy tasks can simplify assessments. However, this complexity may deter new contributors from engaging in open source projects due to the rigorous evaluation processes. Consequently, such barriers can stifle innovation and lead to a less diverse community, highlighting the need to balance capability measurement with fostering participation in open source development.

Shutting down open source initiatives would significantly hinder American technological innovation by stifling collaboration and limiting access to tools and resources. Unfortunately, this may happen under Practice 7.1, Recommendation 1, which encourages developers to "**Share the methodology and results of pre-deployment evaluations of model capabilities, risks, and mitigations, including as much detail about the data and evaluation methodology as can be disclosed without introducing risks to public safety.**"[6] The emphasis on disclosing detailed methodologies and risk management

[2] *What are foundation models?*, Amazon (2024). https://aws.amazon.com/what-is/foundation-models/
[3] *Introducing the Next Generation of Claude*, Anthropic (Mar. 2024). https://www.anthropic.com/news/claude-3-family
[4] Gayathri Siva, *BERT – Bidirectional Encoder Representations from Transformers*, Medium (Nov. 2021). https://gayathri-siva.medium.com/bert-bidirectional-encoder-representations-from-transformer-8c84bd4c9021
[5] NIST p.10, *Managing Misuse Risk for Dual-Use Foundation Models*, Dept. of Commerce (July 2024). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-1.ipd.pdf
[6] NIST p.16, *Managing Misuse Risk for Dual-Use Foundation Models*, Dept. of Commerce (July 2024). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-1.ipd.pdf

processes may deter participation in open source initiatives due to concerns about exposing sensitive information and potential liabilities.

Open source projects fuel creativity and advancement by allowing developers and researchers to build on each other's work, leading to faster breakthroughs and more diverse technological solutions. Additionally, by leveraging foundation models, startups can accelerate their development processes, reduce costs, and bring new products and services to market. Without open access, innovation would slow, as smaller companies, startups, and independent developers would face increased barriers to entry, reducing competition and potentially allowing other countries to outpace the US in tech advancements.

Concerns about the harms of foundation models remain largely theoretical. The real-world benefits of open source innovation far outweigh hypothetical risks, which can be managed through responsible practices.

Voluntary and responsible practices help in measuring misuse risks
Additionally, as stated in Practice 4.1, accurate risk measure is critical. However, in doing so, we should consider the marginal risk of newer technologies, like models, with what exists to make a meaningful risk determination. Suppose a new AI technology introduces a marginally higher risk than existing systems. In that case, protocol should focus on targeted measures that address these specific risks rather than imposing broad regulations that could hinder the overall competitive landscape. For example, to the extent that generative AI models may pose additional risks related to information integrity, we should focus on interventions that address that directly. We applaud NIST for acknowledging that "**misuse risks are not a product of the model alone – they result in part from malicious actors' motivations, resources, constraints, as well as society's defense measures against that harm**."[7] We agree that the focus should be addressing the abuse of models rather than limiting the potential benefits of the models themselves.

We commend the AI Safety Institute for taking the initiative to establish voluntary guidelines, reflecting a commitment to responsible practices. However, it is crucial to ensure that foundation models receive robust support, as they are the backbone of future technological advancements. To maintain its competitive edge, the US must prioritize maintaining its position as a global leader in innovation, fostering an environment where cutting-edge technologies can thrive and drive economic growth.

Sincerely,

[7] NIST Introduction, *Managing Misuse Risk for Dual-Use Foundation Models*, Dept. of Commerce (July 2024). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.800-1.ipd.pdf

Todd O’Boyle
Senior Director, Technology Policy
Chamber of Progress