BEFORE THE
NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION

In the Matter of **Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights**	Docket Nos. NTIA-2023-0009; 240216-0052

COMMENTS OF THE CONSUMER TECHNOLOGY ASSOCIATION IN RESPONSE TO THE NTIA'S REQUEST FOR COMMENTS ON DUAL USE FOUNDATION ARTIFICIAL INTELLIGENCE MODELS WITH WIDELY AVAILABLE MODEL WEIGHTS

The Consumer Technology Association® ("CTA") submits this response to the National Telecommunications and Information Administration ("NTIA") Request for Comment ("Request for Comment") on dual use foundation artificial intelligence models with widely available model weights. CTA is North America's largest technology trade association. Our members are the world's leading innovators—from startups to global brands—helping support more than 18 million American jobs. CTA owns and produces CES®—the most influential tech event in the world.

CTA urges the NTIA to proceed with caution when considering whether, or what, new rules may be necessary as the agency explores the benefits, risks and related issues surrounding the development and deployment of artificial intelligence models with widely available model weights ("open weight models"). Artificial Intelligence ("AI") as a category of technologies is not new, but generative AI systems and technologies such as machine learning that underlie AI systems are emerging technologies that are evolving rapidly. Indeed, a recent Federal Trade Commission ("FTC") report found that AI is nascent, varied, and not susceptible to one definition.[1] AI systems that are used to inform a broader system, such as machine vision systems used to read stop signs in autonomous vehicles, have very different risk profiles compared to generative AI tools connected to the Internet which interact with users directly. Industry leaders in the development of AI systems, including generative AI systems, have been actively working to ensure their systems comply with existing laws, such as privacy, consumer protection, and anti-discrimination regulations.

[1] *See* Combatting Online Harms Through Innovation, FTC, at 1 (June 16, 2022), available at https://www.ftc.gov/system/files/ftc_gov/pdf/Combatting%20Online%20Harms%20Through%20Innovation%3B%20Federal%20Trade%20Commission%20Report%20to%20Congress.pdf ("Combatting Online Harms Report") ("AI is defined in many ways and often in broad terms. The variations stem in part from whether one sees it as a discipline (e.g., a branch of computer science), a concept (e.g., computers performing tasks in ways that simulate human cognition), a set of infrastructures (e.g., the data and computational power needed to train AI systems), or the resulting applications and tools.").

As such, before supporting or proposing policies that call for new regulations in this area, NTIA should ensure it develops a robust record and undertakes sufficient deliberation and consideration of both the benefits and risks presented by the use of AI systems and technology. Any new rules recommended by NTIA should be part of a risk-based, flexible approach that accounts for different use cases and is narrowly tailored to avoid imposing undue burdens on innovation.

In these comments, CTA outlines several factors NTIA should consider as it collects information regarding risks and benefits of open weight models. In these comments, CTA offers its perspective on how best to define and classify the many distinct types of open weight models. These comments also outline numerous benefits associated with open weight models and offer perspectives on how to consider such benefits against marginal risks arising from the use of such models. Finally, CTA demonstrates how industry and civil society continue to work towards developing tools and standards for evaluating these models, and why a light-touch, measured approach to any new rules is critically important to ensuring the United States can benefit from the many gains offered by making open weight models widely available.

I. Efforts to Regulate Emerging AI Models and Technologies Require Due Deliberation and Caution

AI offers tremendous potential for human and societal development: promoting inclusive growth, improving the welfare and well-being of individuals, and enhancing global innovation and productivity. A growing body of research demonstrates AI can identify and mitigate bias in human decision making.[2] Perhaps the leading federal agency focused on AI governance and risk management, the National Institute of Science and Technology ("NIST"), has recently commented that "new AI-enabled systems are revolutionizing and benefitting nearly all aspects of our society and economy – everything from commerce and healthcare to transportation and cybersecurity."[3]

Further, CTA members help promote responsible and trustworthy AI through leadership in the development of emerging practices that mitigate risks, such as the use of federated

[2] *See, e.g.*, Jon Kleinberg et al., *Discrimination in the Age of Algorithms*, 10 J. of Legal Analysis 113, 120 (2019), https://academic.oup.com/jla/article/doi/10.1093/jla/laz001/5476086; Cass R. Sunstein, *Algorithms, Correcting Biases*, 86 Soc. Rsch.: An Int'l Q. 499, 500 (2019), http://eliassi.org/sunstein_2019_algs_correcting_biases.pdf; Kimberly A. Houser, *Can AI Solve the Diversity Problem in the Tech Industry? Mitigating Noise and Bias in Employment Decision-Making*, 22 Stan. Tech. L. Rev. 290, 352 (2019), https://www-cdn.law.stanford.edu/wp-content/uploads/2019/08/Houser_20190830_test.pdf.

[3] Moreover, NIST recognizes that AI "is rapidly transforming our world. Remarkable surges in AI capabilities have led to a wide range of innovations including autonomous vehicles and connected Internet of Things devices in our homes. AI is even contributing to the development of a brain-controlled robotic arm that can help a paralyzed person feel again through complex direct human-brain interfaces." *Artificial Intelligence*, NIST, https://www.nist.gov/artificial-intelligence (last visited Oct. 1, 2022). *See also About Artificial Intelligence*, National Artificial Intelligence Initiative Office, https://www.ai.gov/about/ (last visited Oct. 1, 2022) (explaining that investments in AI technology "have led to transformative advances now impacting our everyday lives, including mapping technologies, voice-assisted smart phones, handwriting recognition for mail delivery, financial trading, smart logistics, spam filtering, language translation, and more. AI advances are also providing great benefits to our social wellbeing in areas such as precision medicine, environmental sustainability, education, and public welfare.").

learning, a machine learning ("ML") approach that learns from a user's interaction with a given device while keeping all the training data on the device, so that the data does not need to be shared with a server.[4] Indeed, CTA has supported efforts at the federal level to develop voluntary risk-based frameworks to address potential AI risks, while enabling stakeholders to maximize the benefits of this technology,[5] an approach reflected in NIST's AI Risk Management Framework ("RMF"), a flexible and voluntary framework for managing AI risks. CTA also has produced consensus standards supporting responsible and trustworthy AI.[6]

NIST's findings and decision to use a voluntary framework suggest it may be premature for NTIA to move forward with broad restrictions on a nascent technology which offers the potential to dramatically improve consumer well-being. This is especially true given public and private sector efforts to establish voluntary risk management frameworks that are tailored to potential risks while still allowing AI to be deployed in beneficial ways. Given the increased use of these voluntary risk management frameworks and the fast-moving pace of development of this technology, NTIA should proceed with caution and avoid adopting overly prescriptive rules. For the same reason, the National Security Commission on Artificial Intelligence's Final Report did not recommend regulation for AI technologies due, in part, to the "speed of technology development by the private sector …"[7] Prescriptive rules would undermine the important work that has been done across the public and private sectors to focus on risk-based approaches. These findings counsel against the adoption of broad prescriptive rules at this time.

II. Response to Certain Questions Framed by NTIA's Request For Comments

A. How should NTIA define "open" or "widely available" when thinking about foundation models and model weights?

Properly defining and scoping models with widely available open weights ("open weight models") is critical to ensure that policy findings and actions are focused and precise. Ill-defined terms and concepts may lead to rules or recommendations that are overbroad and which would likely lead to costly new obligations on model developers or deployers at the risk of undermining innovation. As a foundational principle, NTIA should avoid a binary approach, defining models as either open or closed. In fact, there are gradients of openness in these models. In addition, various AI system components, such as algorithms and training data, can also be made more or less open. Thus, NTIA should avoid a narrow or prescriptive definition to account for the

[4] For example, Google recently published research on Entities as Experts AI, explaining how these systems are answering text-based questions with less data.[4] Google has also published guidance for regulators on how to most effectively regulate AI in its *Recommendations for Regulating AI* paper. Recommendations for Regulating AI, Google, https://ai.google/static/documents/recommendations-for-regulating-ai.pdf (last visited Mar. 24, 2024).

[5] *See, e.g.*, Consumer Technology Association Comments, RFI - NIST AI Risk Management Framework, Docket No. 21076-01510 (filed Sept. 15, 2021), available at https://www.nist.gov/system/files/documents/2021/09/16/ai-rmf-rfi-0087.pdf.

[6] *See*, for example, *The Use of Artificial Intelligence in Health Care: Trustworthiness (ANSI/CTA-2090) (rel. Feb. 2021), available at: https://shop.cta.tech/products/the-use-of-artificial-intelligence-in-healthcare-trustworthiness-cta-2090; and Guidelines for Developing Trustworthy Artificial Intelligence Systems (ANSI/CTA-2096) (rel. Nov. 2021), available at:* https://shop.cta.tech/collections/standards/products/guidelines-for-developing-trustworthy-artificial-intelligence-systems-ansi-cta-2096.

[7] *See* Final Report, National Security Commission on Artificial Intelligence, at 449 (Mar. 19, 2021), *available at* https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.

variability of distinct types of models.[8]

Although openness in models and components vary, a recent study by AI researchers[9] notes that open models share common attributes: (i) *broad access* - open models generally make model weights widely available to the public; (ii) *greater customizability* – these models can be customized for various downstream applications; (iii) *on device capabilities* – open models can directly deployed on local hardware; (iv) *inability to rescind* – access and use rights to these models are not easily revoked by the model developer; and, (v) *limits on downstream moderation* – because inferences can occur on several platforms other than the developer's, monitoring or moderating downstream use is challenging.[10] NTIA should consider these aspects when defining the attributes of openness and thinking about foundation models and model weights.

At the same time, NTIA should avoid conflating open weight models generally with the "dual use foundation models with widely available weights" identified in the recent Executive Order.[11] The E.O. clearly defines the scope of "dual use foundation models with widely available weights" in terms of those models presenting specific risks associated with threats arising from chemical, biological, radiological, or nuclear weapons, cyber-attacks or related risks.[12] However, NTIA must recognize that "foundation models with widely available weights" or "open foundation models" as used in the Request for Comments are <u>not</u> necessarily dual use models. Many "open foundation models" do not present the kinds of risks identified in the E.O., nor do all open foundation models offer dual use capabilities.

NTIA should ensure it adheres to the distinction between models which clearly are within scope of the E.O. defined term ("dual use foundation models with widely available weights") and those that are not. Otherwise, all open foundation models will be treated the same from a risk perspective and will face the attendant obligations of mitigating such risks (i.e. greater regulatory duties). Such an approach would severely limit the use and benefits of such open models and have a significant negative impact on innovation, competitiveness and research opportunities made available through access to open models.

Finally, NTIA should avoid using model computing power thresholds (e.g., FLOPs) as a proxy for classifying open weight models. Instead, any classification should focus on the capabilities of models using <u>qualitative</u> criteria such as those put forward by organizations focused on defining such criteria by leveraging technical expertise, testing, evaluation, and verification of AI systems, such as the AI Verify Foundation.

B. <u>What are the benefits of foundation models with model weights that are widely available as compared to fully closed models?</u>

---

[8] In addition, NTIA should consider harmonizing terminology in this area, such as by defining the spectrum of open models as "open innovation" models.
[9] *On the Societal Impact of Open Foundation Models*, Sayash Kapoor, Rishi Bommasani, et al. (rel. Feb. 27, 2024) (hereafter *Kapoor, et al., Open Foundation Models*) available at: https://arxiv.org/abs/2403.07918.
[10] *Id*. at 3.
[11] Executive Order 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 FR 75191 (rel. Oct. 30, 2023).
[12] *Id*. at Sec. 3(k).

Open weight models offer numerous benefits not available through closed models. Open weight models provide opportunities for increased innovation by enabling broader access and greater customizability to the output behaviors of those models. Because open weight models can be more deeply customized, they can support innovation across a range of applications.[13] Further, models with open weights allow application developers and researchers to perform inference and adaptation locally, which enables adaptation or fine-tuning of models on large proprietary datasets without increasing risks of privacy or data protection.[14]

Further, open weight models enable greater transparency that is considerably higher than their closed counterparts. Indeed, some researchers have recently concluded that "[w]idely available model weights enable external researchers, auditors, and journalists to investigate and scrutinize foundation models more deeply."[15] Greater transparency is likely to reduce potential harms caused by opaque systems and can better enable research on the effects of such models. Many platforms have the ability to aggregate and track portions of open foundation models, which can help visibility and facilitate oversight.[16]

Another benefit of open weight models is that such models lower barriers to entry and thereby help distribute opportunity and mitigate market concentration. Because these models have lower barriers to entry (e.g., cost, expertise), they are more accessible to the general public. Leveraging input and feedback from the broader AI community of researchers and users can help identify and mitigate bugs, biases, and safety issues that may otherwise go unnoticed, ultimately leading to better performing and safer AI products.[17] This lower barrier to entry can help to drive AI research and development by academics or other subject matter experts, enabling communities with bespoke datasets and unique needs to form around specific platforms or industry sectors.[18]

U.S. development of open weight models can further America's position as a leader in technology and innovation. An open, U.S.-promoted AI ecosystem may help foster multilateral collaboration on AI policy and security issues, as it may create a common and more transparent baseline from which all countries can assess the technology.[19]

Finally, whereas closed system developers have the exclusive ability to control and restrict uses cases they deem unacceptable, users of open weight models may have the ability to make these types of decisions themselves.[20] While restricting use cases can be a benefit against

---

[13] *Kapoor, et al., Open Foundation Models* at 4.
[14] *Id.*
[15] *Id.*
[16] Center for Security and Emerging Technology (CSET), Open Foundation Models: Implications of Contemporary Artificial Intelligence (Mar. 12, 2024) available at: https://cset.georgetown.edu/article/open-foundation-models-implications-of-contemporary-artificial-intelligence/. (hereafter "*CSET, Open Foundation Models*").
[17] Open-Sourcing Highly Capable Foundation Models, Elizabeth Seger, Noemi Dreksler, et al., (rel. Sept. 29, 2023) available at: https://arxiv.org/pdf/2311.09227.pdf.
[18] *CSET, Open Foundation Models* available at: https://cset.georgetown.edu/article/open-foundation-models-implications-of-contemporary-artificial-intelligence/.
[19] *Id.*
[20] *Considerations for Governing Open Foundation Models*, Rishi Bommasani, Sayash Kapoor, et al., Stanford University Human-centered Artificial Intelligence (rel. Dec. 2023) available at https://hai.stanford.edu/sites/default/files/2023-12/Governing-Open-Foundation-Models.pdf.

malicious actors, it also can have the effect of concentrating power with a small number of actors. More actors within the ecosystem will enable increased scrutiny of open models, which can help to crowdsource vulnerability and safety discovery that, in turn, will facilitate the discovery and implementation of patches and safety/security measures.[21]

C. How do the risks associated with making model weights widely available compare to the risks associated with non-public model weights?

NTIA's consideration of risks associated with open weight models should focus on marginal risks arising from such models. As recognized by *Kapoor, et al*., risk analysis in this context must be grounded in a recognition that open models only present additional risks of misuse or harm beyond those harms that already exist via from the use of existing technologies (i.e., "marginal risk").[22] Indeed, while risks associated with access to widely available open weight models should not be dismissed, these concerns should be weighed against the recognition that certain risks apply to all model uses. Similarly, many mitigation tools available for existing risks arising from the use of closed models may be sufficient to guard against risks perceived to be caused or exacerbated by widely available model weights. Thus, NTIA should focus its analysis on marginal risks and within the context of generally available mitigation tools. With this in mind, NTIA should consider risks arising from open models holistically, accounting for only the marginal risk presented by open weight models, balanced against available mitigation tools and the significant benefits arising from the availability of open models.

Further, one of the notable benefits of open models is that they facilitate opportunities for users to develop tailored mitigation tools and techniques to address risks unique to specific uses cases or operational scenarios. The clearest evidence of this benefit is in the area of cybersecurity, where crowdsourced solutions and mitigation techniques are helping to combat the ever evolving cybersecurity risk matrix. Thus, while openness may present certain unique marginal risks, the same attribute offers the benefit of developing mitigation tools from a broader universe of users.

A fundamental risk of open weight models is the inability to rescind model access once a model weight has been made widely available. Although the model developer can limit access upon learning of a new marginal risk, existing copies of the model weights cannot be revoked. Certain risks, such as difficulties in establishing the validity of results or risks caused by bias within the model, may be mitigated by the release of training data or source code associated with fine tuning, pretraining, or deploying a model that is simultaneously widely available. These risks also must be balanced against the potential benefits of open model weights, which include the democratizing of AI by making it more widely accessible to more people in more parts of the world; driving AI use case innovation, experimentation, and adoption; and supporting AI safety research.

With respect to privacy risks in particular, the openness of a system allows for innovation in the privacy and security domains in ways that closed systems would not because they allow

[21] *CSET, Open Foundation Models* available at: https://cset.georgetown.edu/article/open-foundation-models-implications-of-contemporary-artificial-intelligence/.
[22] *Kapoor, et al., Open Foundation Models,* at 5-8.

good actors, such as cybersecurity and privacy researchers, to use the fine-tuning options of widely available model weights to stress-test the existing privacy and security controls of AI systems. Additionally, with limited exception, the privacy risks that could result from widely available model weights would not be a marginal risk because the provision of sensitive data to a third party service provider always presents an inherent risk of exposure to sensitive or confidential data if that third-party actor is compromised.[23] This type of risk can be mitigated by organizational policies that limit the input of sensitive data for fine-tuning. Where this would not be possible, existing cybersecurity tools and frameworks can likely be leveraged to meet many of the needs for protecting the security and privacy of AI systems with open model weights.

Another risk mitigation benefit unique to open weight models is that entities using such models are often able to deploy the model locally on a device or at the network edge. This generally reduces privacy risks by allowing the entity to use the model without the need to share sensitive or confidential business data with third parties.[24] This benefit is particularly important in domains involving significant amounts of PII, such as healthcare and financial services.

Accordingly, NTIA's analysis of risks in this area should focus on whether there are marginal risks posed by a model, whether such marginal risks are significant or not, and then consider whether such risks are outweighed by the benefits of widely available model weights in this circumstance, or whether availability to the model weights should be restricted based on the level of anticipated marginal risk.[25]

D. What are current or potential voluntary, domestic regulatory, and international mechanisms to manage the risks and maximize the benefits of foundation models with widely available weights?

Risk management protocols and procedures are beginning to emerge, but continued development and refinement of benchmarks, evaluations and assessments is necessary. This work is critically important and requires a precise articulation of the actual, rather than speculative, marginal risks presented in this area.

Entities leading in the development of the risk management mechanisms are numerous and include most notably NIST, through its work on a risk management framework and related questions. In addition, standards development organizations such as the International Organization for Standardization (ISO), CTA and others, plus numerous industry-led initiatives such as the AI Alliance, the Partnership on AI, and the Frontier Model Forum, along with academic initiatives such as Stanford University's Human-Centered Artificial Intelligence and other non-governmental bodies, play a critical role in developing technical standards. Sector specific groups also can be involved in defining future AI governance. Finally, open-source model evaluation tools are available in the market today[26] and those offerings are expected to

[23] *See* Privacy Risks of LLM Fine Tuning, Daniel Huynh (rel. Nov. 22, 2023), available at: https://blog.mithrilsecurity.io/privacy-risks-of-llm-fine-tuning/.
[24] *Kapoor, et al., Open Foundation Models* at 3.
[25] Of course, not all downstream uses are foreseeable, which may limit the utility of this approach in certain circumstances.
[26] *See, e.g.*, Tensorflow Model Remediation, https://github.com/tensorflow/model-remediation (library that provides solutions for practitioners working to create and train models in a way that reduces or eliminates user harm resulting

continue to be available as this technology matures.

NTIA should leverage the important work of these organizations and permit industry and civil society to continue to define appropriate benchmarks, evaluations, and assessment criteria independently, especially in the area of technical standards. Beyond these areas, NTIA's focus should be on broader structural questions around regulation, liability and competition. For example, distinguishing between the role of developer and deployer in this ecosystem is critical. Consistent with the CTA's National AI Policy and Regulatory Framework, because developers of open weight models have limited control over downstream uses, deployers should be responsible and accountable for potential risks or harms that occur when the open weight models are deployed.[27]

E. In the face of continually changing technology, and given unforeseen risks and benefits, how can governments, companies, and individuals make decisions or plans today about open foundation models that will be useful in the future? How should these potentially competing interests of innovation, competition, and security be addressed or balanced?

AI provides the U.S. with the opportunity to advance innovation, boost economic productivity, and maintain global leadership in an emerging technology critical to American security interests against adversaries.[28] Any policy development or governance frameworks that focus on hypothetical concerns that AI could surpass human abilities could put in jeopardy the opportunity to address more near-term governance, safety and security concerns, and have the potential to suppress innovation.[29] It is increasingly important that such concerns do not exclude important perspectives from stakeholders in such a manner as to erode public trust in AI policy development institutions.[30]

A comprehensive approach would involve the appropriate release of models with some level of openness when the benefits significantly exceed the risks. In addition to the decision to make a model available, there are complementary mechanisms for managing risk such as investment in broad-based innovation and safety research including partnerships between industry organizations, such as the Partnership on AI and Frontier Model Forum, and stress testing from external partners. Because developers have no control over final end uses of open foundation models, it is critically important that liability for this balance lies at the closest point

---

from underlying performance biases); Fairness Comparison, https://github.com/algofairness/fairness-comparison (facilitates benchmarking of fairness aware machine learning algorithms).

[27] CTA, National AI Policy and Regulatory Framework (re. Sep. 23, 2023), *available at: https://cdn.cta.tech/cta/media/media/pdfs/ai-policy.pdf?_gl=1*1pthhjn*_ga*MjAyNDE5MTA2Mi4xNzExNTgyMzgw*_ga_5P7N8TBME7*MTcxMTU4MjM3OS4xLjEuMTcxMTU4MzIyOC42MC4wLjA.*

[28] AEI, AI Is a National Security Lifeline, by Klon Kitchen (rel. Aug. 15, 2023), available at: https://www.aei.org/foreign-and-defense-policy/ai-is-a-national-security-lifeline/.

[29] Center for Security and Emerging Technology, Commentary: Balancing AI Governance with Opportunity (rel. Nov. 30, 2023) available at: https://cset.georgetown.edu/article/commentary-balancing-ai-governance-with-opportunity/.

[30] Carnegie Endowment for International Peace, How Hype Over AI Superintelligence Could Lead Policy Astray (rel. Sept. 14 2023), available at: https://carnegieendowment.org/2023/09/14/how-hype-over-ai-superintelligence-could-lead-policy-astray-pub-90564.

to the end user of an AI product.

Respectfully submitted,

CONSUMER TECHNOLOGY ASSOCIATION

/s/ Douglas K. Johnson
Douglas K. Johnson
Vice President, Emerging Technology Policy

/s/ Michael Petricone
Michael Petricone
Sr. Vice President, Government and Regulatory Affairs

Dated: March 27, 2024