National Telecommunications and Information Administration Request for Comments on Dual Use Foundation Models and Widely Available Model Weights
March 26, 2024

BSA appreciates the opportunity to provide comments on the National Telecommunications and Information Administration's (NTIA) Request for Comments (RFC) on Dual Use Foundation Models and Widely Available Model Weights.

BSA is the leading advocate for the global software industry.[1] BSA members are at the forefront of developing cutting-edge services — including AI — and their products are used by businesses across every sector of the economy.[2] For example, BSA members provide tools including cloud storage and data processing services, customer relationship management software, human resource management programs, identity management services, cybersecurity services, and collaboration software. BSA members are on the leading edge of providing AI-enabled products and services. As a result, they have unique insights into the technology's tremendous potential to spur digital transformation and the policies that can best support the responsible use of AI.

BSA's views are informed by our experience working with member companies to develop the BSA Framework to Build Trust in AI,[3] a risk management framework we published almost three years ago to help companies mitigate the potential for unintended bias in AI systems. Built on a vast body of research and informed by the experience of leading AI developers, the BSA Framework outlines a lifecycle-based approach for performing impact assessments and highlights corresponding best practices.[4] Our experience on these issues informs our recommendations below.

[1] BSA's members include: Adobe, Alteryx, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, CNC/Mastercam, Databricks, DocuSign, Dropbox, Elastic, Graphisoft, Hubspot, IBM, Informatica, Kyndryl, MathWorks, Microsoft, Okta, Oracle, PagerDuty, Palo Alto Networks, Prokon, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Workday, Zendesk, and Zoom Video Communications, Inc.

[2] *See* BSA | The Software Alliance, Artificial Intelligence in Every Sector, *available at* https://www.bsa.org/files/policy-filings/06132022bsaaieverysector.pdf.

[3] *See* BSA | The Software Alliance, Confronting Bias: BSA's Framework to Build Trust in AI, *available at* https://www.bsa.org/reports/confronting-bias-bsas-framework-to-build-trust-in-ai.

[4] BSA has testified before the United States Congress and the European Parliament on the Framework and its approach to mitigating AI-related risks*. See, e.g.,* Testimony of Victoria Espinel, Public Hearing on AI & Bias, Special Committee on Artificial Intelligence in a Digital Age, European Parliament, Nov. 30, 2021, *available at* https://www.europarl.europa.eu/cmsdata/244265/AIDA_Verbatim_30_November_2021_EN.pdf; Testimony of Victoria Espinel, The Need for Transparency in Artificial Intelligence, Before the Senate Committee on Commerce, Science, and Transportation Subcommittee on Consumer Protection, Product Safety, and Data Security, *available at* https://www.bsa.org/files/policy-filings/09122023aitestimonyoral.pdf.

AI systems with widely available model weights are an increasingly important part of the AI supply chain. Many companies increasingly incorporate ‘fine-tuned’ (retrained) open large language models (LLMs) in a wide variety of business applications. Proprietary foundation models also continue to spur innovation, and the US government should support both distribution models.

To the degree that AI creates risks that companies should guard against, many of these risks are attributable to AI models broadly, rather than open foundation models specifically. Because the development of open foundation models does not occur in the United States alone, restrictions on open foundation models could cut off regulators’ access to essential information about how systems operate in practice and ultimately fail to mitigate potential risks.

The RFC asks about a range of issues related to open foundation models, including their benefits, their risks, potential safeguards, and the role of government in imposing restrictions on the availability of model weights. We address these issues and recommend the US government:

- Recognize the substantial benefits that open foundation models provide to both consumers and businesses;
- Avoid restricting the availability of open foundation models;
- Ground policies that address risks of open foundation models on empirical evidence; and
- Encourage the implementation of safeguards to enhance the safety of open foundation models.

I. Open foundation models provide substantial benefits to businesses and consumers.

AI provides immeasurable economic and societal benefits, and open foundation models play an important role in fueling this innovation. Indeed, the benefits that open foundation models provide are substantial and include:

- **Lower Cost, More Competition, and Democratization of AI.** LLMs are expensive to operate, often 10 cents or more per query, and using an open LLM brings the operating cost down considerably. Open foundation models also lower barriers to entry, improve competition, decrease the cost of ownership, and increase the range of alternative options, thus democratizing the use and control of AI. Availability of open foundation models also enables companies that don’t have access to significant levels of compute to bridge the gap and offer foundation model capabilities that match the outputs of proprietary models. For example, a company may start with an open foundation model, fine tune that model, and integrate the revised model into the products and services it provides to consumers and businesses. Open AI models also reduce vendor lock-in, giving users greater flexibility and ability to control their AI system.

- **Increased Transparency and Quality Control.** Open foundation models allow organizations to understand exactly what is happening in AI models because the source code — and typically the model weights and training data — are freely available for access and inspection to any third party that may wish to analyze them, including regulators and researchers. This accessibility and transparency can enable rapid identification and mitigation of problems that may arise, permit greater understanding of how an AI system works, and generally enhance trust in AI.

- **Safety and Resiliency.** Access to open foundation models is a driving force behind research on increasing safety and reducing AI risks, including alignment methods. Researchers and other experts use open foundation models to develop solutions to a range of issues relevant to both proprietary and open foundation models, including interpretability, security, and safety. A decentralized open AI ecosystem also increases the availability of different foundation models, which enhances resiliency because there is no single point of failure in the event that a particular model becomes impaired or unavailable. Indeed, many enterprise firms leverage multiple open and proprietary models to power products and features within their ecosystems precisely as a means of ensuring resiliency. The availability of open AI models can make the AI ecosystem more, not less, safe.

- **Increased Customization.** With open foundation models, organizations have an opportunity to create customized applications by fine-tuning or otherwise adapting the model for particular purposes, which expands the number of potential applications and increases innovation. For example, enterprise customers of video communications providers can, within their own closed data ecosystem, use open LLMs to power features that provide tailored summaries of online meetings. In addition to customizing how the model operates, customizable model sizing also provides companies with an alternative to operating a larger model through an API where a smaller model may be perfectly capable for a given use case. In addition to making available AI tools to companies with less computational power, using smaller models helps companies of all sizes decrease compute workloads and impact on the environment.

- **Advanced Scientific Research.** The broader access provided by open foundation models accelerates scientific research in a range of fields, and it is essential for reproducibility of research. Notably, open AI models have enhanced healthcare research in medical imaging analysis, which could lead to more accurate diagnoses and better treatment decisions.

- **Promoting Equity and Inclusion.** The availability of open foundation models increases access to cutting-edge AI tools. As a result, it lowers barriers to obtaining skills that are necessary to have careers in AI, expanding opportunities for diverse communities.

II. The US government should avoid restricting the availability of open foundation models.

The RFC solicits information on the role the US government should play in supporting or restricting the availability of model weights, and on legal or other measures that could be employed to prevent widespread availability of open foundation model weights. We recommend the US government not restrict the availability of open foundation models. Instead, it should support the further development of a robust AI ecosystem, including open foundation models. Any specific policy options for open foundation models should be considered only as any marginal risk posed by such models are better understood.

Restricting foundation models with widely available model weights would significantly curtail the benefits of open foundation models. Restrictions are also unlikely to mitigate perceived risks, because the development of open foundation models does not occur in the United

States alone. Indeed, such restrictions may instead cut off an essential pipeline of information for researchers, including those focused on advancing AI safety solutions. Restricting the availability of open foundation models in the United States could also frustrate the US government's regulatory efforts by limiting access to key information that provides an understanding of how particular AI models operate. Internationally, the development and use of open foundation models would continue to be available to global partners and adversaries, while US organizations would be stymied, further hampering US economic and security interests.

The RFC also inquires about circumstances in which the government contracts with companies that use open foundation models. The Administration should not limit the government's ability to contract with companies that use open foundation models, nor should it require disclosures regarding the use of these models. Open-source components are ubiquitous in software and are a key part of the technological landscape that fuels innovation. Companies should be incentivized, not discouraged, from leveraging these resources to expand the diversity and capability of available applications. The government should also benefit from the innovation created through open foundation models, which aligns with the Administration's IT modernization goals.

Finally, the US government should support a robust AI ecosystem with open models that is government-wide and does not vary by sector. The US AI Safety Institute could be particularly helpful in advancing the responsible development of open foundation models, including supporting more research on credible risks and safety safeguards. Global interoperability is also critical to effective AI policies, and the United States should work with its allies to ensure that the global policy landscape promotes the development and use of responsible AI, including open foundation models.

III. Policies aimed at addressing the risks of open foundation models should be grounded in empirical evidence.

AI provides immense benefits to consumers and businesses, including stimulating economic growth and solving complex societal challenges. In some contexts, however, certain uses of AI can exacerbate existing risks, including risks of biased outputs and disinformation. BSA members are committed to taking steps to mitigate AI risks as they develop and use AI responsibly. That is why BSA worked with member companies to develop the BSA Framework to Build Trust in AI, which was released in 2021 and is designed to help organizations mitigate the potential for unintended bias in AI systems by adopting a lifecycle-based approach to risk management.

The RFC acknowledges that open foundation models create benefits, but it also indicates that open foundation models could "engender substantial harm," such as risks to security and equity. Although we agree that AI has the potential to create risks that companies should guard against, there are important questions about how much such risks can be attributed to the availability of model weights. Leading researchers from Stanford, Princeton, and other notable organizations recently released research highlighting commonly cited risks of open foundation models, including biosecurity and cybersecurity.[5] They emphasized that other types of technology already present similar risks, and that more empirical evidence is necessary to quantify the marginal risk created by open foundation models — i.e., the amount of new risk that goes beyond those already posed by

[5] Sayash Kapoor et al., On the Societal Impact of Open Foundation Models, Feb. 27, 2024, at 2, *available at* https://crfm.stanford.edu/open-fms/paper.pdf.

proprietary foundation models or other pre-existing technologies.[6] As one example, the paper highlights concerns about open foundation models generating accurate information about pandemic-causing pathogens — and notes that similar information is available on public internet search engines.[7]

As the US government develops policies governing open foundation models, it should ground those policies in credible evidence of the incremental risks presented by the availability of open models as compared to proprietary models or other technologies, like search engines. To the extent that evidence remains unavailable, the government should take care to avoid premature action. Even in areas where more impact of open foundation models has been observed — such as facilitating disinformation — the US government should weigh these outcomes against the considerable benefits that open foundation models provide and create more targeted policy solutions that address actual harms. The government also may want to consider ways to encourage solutions that use models with widely available model weights to address such risks, to avoid diminishing the benefits of open models in other contexts.

IV. AI developers and deployers should be encouraged to implement safeguards to advance responsible AI in both open and proprietary AI systems.

There are a range of safeguards that developers and deployers of AI can implement to identify, mitigate, and ultimately reduce potential risks associated with AI systems. We focus on five particularly important safeguards: risk management programs, impact assessments, information-sharing, model evaluations, and safety measures. In many cases, these safeguards can help to identify and address risks across AI systems, including for both open and proprietary AI systems.

Risk management programs. Companies can adopt risk management programs to identify the personnel, policies, and processes necessary to manage AI risks. A strong risk management program can benefit companies that develop or use either proprietary or open AI systems; companies that use both open and proprietary systems can also benefit from a risk management program that helps them holistically identify risks across multiple AI systems. To implement a risk management program, a company can adopt a range of important corporate governance elements, including clearly assigning roles and responsibilities to key personnel, establishing formal policies on their development and/or use of AI, identifying their evaluation mechanisms, ensuring executive oversight, performing impact assessments for high-risk AI, and creating internal independent review mechanisms, such as interdepartmental governance or ethics committees, to evaluate and address AI issues that pose high risks. These steps align with the AI Risk Management Framework (RMF) developed by the National Institute of Standards and Technology (NIST).

Impact assessments. Impact assessments are important accountability tools that help developers and deployers identify and mitigate risks associated with open or proprietary high-risk AI systems. Impact assessments should focus on high-risk AI systems, to ensure that organizations devote resources to addressing systems that pose the greatest potential risks. Importantly, there is no "one-size-fits-all" approach to evaluating and mitigating risks of AI; impact assessments should be tailored to address the nature of the system at issue, the type of harms the system may pose, and the role of the actor along the AI value chain. For example, a company developing an AI system based on an open foundation model may

[6] *Id.* at 2, 8.

[7] *Id.*

be well situated to assess whether additional safeguards are necessary to protect privacy or security in the context of that system, determine performance metrics, and improve representativeness of the data used to fine-tune the model.

Information sharing. Developers of foundation models should ensure that they provide to downstream providers or otherwise make available transparent information documenting key aspects of the model, including its design features, capabilities, a summary of the type of data used in training, known limitations, and factors relating to safety and security features. This information is consistent with model cards provided for some existing open AI models and will be helpful to organizations fine-tuning the model and integrating it into other products and services.

Model evaluations. Model evaluations can be an important mechanism for surfacing problems with an AI model. Testing is a key component of these evaluations and can identify safety and bias issues. Techniques can include adversarial testing (i.e., red-teaming) and vulnerability scanning. Cross-disciplinary review will also be helpful to detect and address a wider array of issues that may arise.

Safety. Developers of open foundation models should diligently follow industry standard safety procedures in developing open foundation models. There are a range of options available, and research continues to explore new technical mitigation strategies.

These safeguards can be applied in different ways by different companies, depending on their role in developing or deploying an AI system. For example, the developer of an open foundation model may adopt a risk management program, to ensure that the company has personnel and policies that govern its development of the AI model. A company developing its own AI systems based on the open foundation model may focus on additional safeguards appropriate to the context in which that system will be used, such as conducting an impact assessment that identifies the specific risks likely to arise from using that AI system in the company's products and services. A deployer may also conduct an impact assessment, focusing on the risks presented by its context of use, and create feedback mechanisms for addressing issues that arise after deployment.

* * *

Thank you for the opportunity to provide comments. We look forward to serving as a resource as you continue to consider AI policy issues.

Respectfully submitted,

Shaundra Watson

Shaundra Watson
Senior Director, Policy
BSA | The Software Alliance