acceptable estimate of the individual marine mammals taken is available, if the estimated number of individual animals taken is up to, but not greater than, one-third of the best available abundance estimate, NMFS will determine that the numbers of marine mammals taken of a species or stock are small. For more information please see NMFS' discussion of the MMPA's small numbers requirement provided in the final rule (86 FR 5322, 86 FR 5438, January 19, 2021).

The take numbers for authorization, which are determined as described above, are used by NMFS in making the necessary small numbers determinations through comparison with the best available abundance estimates (see discussion at 86 FR 5322, 86 FR 5391, January 19, 2021). For this comparison, NMFS' approach is to use the maximum theoretical population, determined through review of current stock assessment reports (SAR; *https:// www.fisheries.noaa.gov/national/ marine-mammal-protection/marine-mammal-stock-assessments*) and model-predicted abundance information

(*https://seamap.env.duke.edu/models/ Duke/GOM/*). For the latter, for taxa where a density surface model could be produced, we use the maximum mean seasonal (*i.e.,* 3-month) abundance prediction for purposes of comparison as a precautionary smoothing of month-to-month fluctuations and in consideration of a corresponding lack of data in the literature regarding seasonal distribution of marine mammals in the GOM. Information supporting the small numbers determinations is provided in table 1.

TABLE 1—TAKE ANALYSIS

Species	Authorized take [1]	Abundance [2]	Percent abundance
Rice's whale	0	51	n/a
Sperm whale	132	2,207	6.0
Kogia spp	[3]50	4,373	1.1
Beaked whales	580	3,768	15.4
Rough-toothed dolphin	100	4,853	2.1
Bottlenose dolphin	473	176,108	0.3
Clymene dolphin	281	11,895	2.4
Atlantic spotted dolphin	189	74,785	0.3
Pantropical spotted dolphin	1,274	102,361	1.2
Spinner dolphin	341	25,114	1.4
Striped dolphin	110	5,229	2.1
Fraser's dolphin	[4]32	1,665	1.9
Risso's dolphin	83	3,764	2.2
Melon-headed whale	185	7,003	2.6
Pygmy killer whale	43	2,126	2.0
False killer whale	69	3,204	2.2
Killer whale	0	267	n/a
Short-finned pilot whale	53	1,981	2.7

[1] Scalar ratios were not applied in this case due to brief survey duration.
[2] Best abundance estimate. For most taxa, the best abundance estimate for purposes of comparison with take estimates is considered here to be the model-predicted abundance (Roberts *et al.,* 2016). For those taxa where a density surface model predicting abundance by month was produced, the maximum mean seasonal abundance was used. For those taxa where abundance is not predicted by month, only mean annual abundance is available. For Rice's whale and killer whale, the larger estimated SAR abundance estimate is used.
[3] Includes 3 take by Level A harassment and 47 takes by Level B harassment.
[4] Modeled exposure estimate less than assumed average group size (Maze-Foley and Mullin, 2006).

Based on the analysis contained herein of LLOG's planned survey activity described in its LOA application and the anticipated take of marine mammals, NMFS finds that small numbers of marine mammals will be taken relative to the affected species or stock sizes (*i.e.,* less than one-third of the best available abundance estimate) and therefore the taking is of no more than small numbers.

Authorization

NMFS has determined that the level of taking for this LOA request is consistent with the findings made for the total taking allowable under the incidental take regulations and that the amount of take authorized under the LOA is of no more than small numbers. Accordingly, we have issued an LOA to LLOG authorizing the take of marine mammals incidental to its geophysical survey activity, as described above.

Dated: February 20, 2024.

Kimberly Damon-Randall,

Director, Office of Protected Resources, National Marine Fisheries Service.

[FR Doc. 2024–03788 Filed 2–23–24; 8:45 am]

BILLING CODE 3510–22–P

DEPARTMENT OF COMMERCE

National Telecommunications and Information Administration

[Docket No. 240216–0052]

RIN 0660–XC060

Dual Use Foundation Artificial Intelligence Models With Widely Available Model Weights

AGENCY: National Telecommunications and Information Administration, Department of Commerce.

ACTION: Notice; request for comment.

SUMMARY: On October 30, 2023, President Biden issued an Executive order on ''Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,'' which directed the Secretary of Commerce, acting through the Assistant Secretary of Commerce for Communications and Information, and in consultation with the Secretary of State, to conduct a public consultation process and issue a report on the potential risks, benefits, other implications, and appropriate policy and regulatory approaches to dual-use foundation models for which the model weights are widely available. Pursuant to that Executive order, the National Telecommunications and Information Administration (NTIA) hereby issues this Request for Comment on these issues. Responses received will be used to submit a report to the President on the potential benefits, risks, and implications of dual-use foundation

models for which the model weights are widely available, as well as policy and regulatory recommendations pertaining to those models.

DATES: Written comments must be received on or before March 27, 2024.

ADDRESSES: All electronic public comments on this action, identified by *Regulations.gov* docket number NTIA–2023–0009, may be submitted through the Federal e-Rulemaking Portal at *https://www.regulations.gov.* The docket established for this request for comment can be found at *www.Regulations.gov,* NTIA–2023–0009. To make a submission, click the "Comment Now!" icon, complete the required fields, and enter or attach your comments. Additional instructions can be found in the "Instructions" section below, after **SUPPLEMENTARY INFORMATION**.

FOR FURTHER INFORMATION CONTACT: Please direct questions regarding this Request for Comment to Travis Hall at *thall@ntia.gov* with "Openness in AI Request for Comment" in the subject line. If submitting comments by U.S. mail, please address questions to Bertram Lee, National Telecommunications and Information Administration, U.S. Department of Commerce, 1401 Constitution Avenue NW, Washington, DC 20230. Questions submitted via telephone should be directed to (202) 482–3522. Please direct media inquiries to NTIA's Office of Public Affairs, telephone: (202) 482–7002; email: *press@ntia.gov.*

SUPPLEMENTARY INFORMATION:

Background and Authority

Artificial intelligence (AI)[1] has had, and will have, a significant effect on society, the economy, and scientific progress. Many of the most prominent models, including the model that powers ChatGPT, are "fully closed" or "highly restricted," with limited or no public access to their inner workings.

The recent introduction of large, publicly-available models, such as those from Google, Meta, Stability AI, Mistral, the Allen Institute for AI, and EleutherAI, however, has fostered an ecosystem of increasingly "open" advanced AI models, allowing developers and others to fine-tune models using widely available computing.[2]

Dual use foundation models with widely available weights (referred to here as open foundation models) could play a key role in fostering growth among less resourced actors, helping to widely share access to AI's benefits.[3] Small businesses, academic institutions, underfunded entrepreneurs, and even legacy businesses have used these models to further innovate, advance scientific knowledge, and gain potential competitive advantages in the marketplace. The concentration of access to foundation models into a small subset of organizations poses the risk of hindering such innovation and advancements, a concern that could be lessened by availability of open foundation models. Open foundation models can be readily adapted and fine-tuned to specific tasks and possibly make it easier for system developers to scrutinize the role foundation models play in larger AI systems, which is important for rights- and safety-impacting AI systems (*e.g.* healthcare, education, housing, criminal justice, online platforms etc.).[4] These open foundation models have the potential to help scientists make new medical discoveries or even make mundane, time-consuming activities more efficient.[5]

Open foundation models have the potential to transform research, both within computer science[6] and through supporting other disciplines such as medicine, pharmaceutical, and scientific research.[7] Historically, widely available programming libraries have given researchers the ability to simultaneously run and understand algorithms created by other programmers. Researchers and journals have supported the movement towards open science,[8] which includes sharing research artifacts like the data and code required to reproduce results.

Open foundation models can allow for more transparency and enable broader access to allow greater oversight by technical experts, researchers, academics, and those from the security community.[9] Foundation models with widely available model weights could also promote competition in downstream markets for which AI models are a critical input, allowing smaller players to add value by adjusting models originally produced by the large developers.[10] The accessibility

[1] Artificial Intelligence (AI) "has the meaning set forth in 15 U.S.C. 9401(3): a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action." *see* Executive Office of the President, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 FR 75191 (November 1, 2023) *https:// www.federalregister.gov/documents/2023/11/01/ 2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.* "AI Model" means "a component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs." *see* Id.

[2] *See e.g.,* Zoe Brammer, How Does Access Impact Risk? Assessing AI Foundation Model Risk Along a Gradient of Access, The Institute for Security and Technology (December 2023) *https://securityand technology.org/wp-content/uploads/2023/12/How-Does-Access-Impact-Risk-Assessing-AI-Foundation-Model-Risk-Along-A-Gradient-of-Access-Dec-2023.pdf;* Irene Solaiman, The Gradient of Generative AI Release: Methods and Considerations, arXiv:2302.04844v1 (February 5, 2023); *https://arxiv.org/pdf/2302.04844.pdf.*

[3] *See e.g.,* Elizabeth Seger et al., Open-Sourcing Highly Capable Foundation Models, Centre for the Governance of AI (2023) *https://cdn.governance.ai/ Open-Sourcing_Highly_Capable_Foundation_ Models_2023_GovAI.pdf.*

[4] *See e.g.,* Executive Office of the President: Office of Management and Budget, Proposed Memorandum For the Heads of Executive Departments and Agencies (November 3, 2023) *https://www.whitehouse.gov/wp-content/uploads/ 2023/11/AI-in-Government-Memo-draft-for-public-review.pdf;* Cui Beilei et al., Surgical-DINO: Adapter Learning of Foundation Model for Depth Estimation in Endoscopic Surgery, arXiv:2401.06013v1 (January 11, 2024) *https://arxiv.org/pdf/ 2401.06013.pdf* (Using low-ranked adaptation, or LoRA, in a foundation model to help with surgical depth estimation for endoscopic surgeries).

[5] *See e.g.,* Shaoting Zhang, On the Challenges and Perspectives of Foundation Models for Medical

Image Analysis, arXiv:2306.05705v2 (November 23, 2023), *https://arxiv.org/pdf/2306.05705.pdf.*

[6] *See e.g.,* David Noever, Can Large Language Models Find And Fix Vulnerable Software?, arxiv 2308.10345 (August 20, 2023) *https://arxiv.org/abs/ 2308.10345;*[6] Andreas Stöckl, Evaluating a Synthetic Image Dataset Generated with Stable Diffusion, Proceedings of Eighth International Congress on Information and Communication Technology Vol. 693 (July 25, 2023) *https:// link.springer.com/chapter/10.1007/978-981-99-3243-6_64.*

[7] *See e.g.,* Kun-Hsing Yu et al., Artificial intelligence in healthcare, Nature Biomedical Engineering Vol. 2 719–731 (October 10, 2018) *https://www.nature.com/articles/s41551-018-0305-z#citeas;* Kevin Maik Jablonka et al., 14 examples of how LLMs can transform materials science and chemistry: a reflection on a large language model hackathon, Digital Discovery 2 (August 8, 2023) *https://pubs.rsc.org/en/content/articlehtml/2023/ dd/d3dd00113j.*

[8] *See e.g.,* Harvey V. Fineberg et al., Consensus Study Report: Reproducibility and Replicability in Science, National Academies of Sciences (May 2019) *https://nap.nationalacademies.org/resource/ 25303/R&R.pdf;* Reporting standards and availability of data, materials, code and protocols, *https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards;* Science, Science Journals: Editorial Policies, *https:// www.science.org/content/page/science-journals-editorial-policies#data-and-code-deposition;* Edward Miguel, Evidence on Research Transparency in Economics, Journal of Economic Perspectives Vol. 35 No. 3 (2021) *https:// www.aeaweb.org/articles?id=10.1257/jep.35.3.193.*

[9] *See e.g.,* Rishi Bommasani et al., Considerations for Governing Open Foundation Models, Stanford University Human-Centered Artificial Intelligence (December 2023) *https://hai.stanford.edu/sites/ default/files/2023-12/Governing-Open-Foundation-Models.pdf.*

[10] *See, e.g.,* Jai Vipra and Anton Korinek, Market concentration implications of foundation models: The Invisible Hand of ChatGPT, Brookings Inst. (2023) *https://www.brookings.edu/articles/market-concentration-implications-of-foundation-models-the-invisible-hand-of-chatgpt/.*

of open foundation models also provides tools for individuals and civil society groups to resist authoritarian regimes, furthering democratic values and U.S. foreign policy goals.

While open foundation models potentially offer significant benefits, they may pose risks as well. Foundation models with widely-available model weights could engender substantial harms, such as risks to security, equity, civil rights, or other harms due to, for instance,[11] affirmative misuse, failures of effective oversight, or lack of clear accountability mechanisms.[12] Others argue that these open foundation models enable development of attacks against proprietary models due to similarities in the data sets used to train them.[13] The wide availability of dual use foundation models with widely available model weights and the continually shrinking amount of compute necessary to fine-tune these models together create opportunities for malicious actors to use such models to engage in harm.[14] The lack of monitoring of open foundation models may worsen existing challenges, for example, by easing creation of synthetic non-consensual intimate images or enabling mass disinformation campaigns.[15]

On October 30, 2023, President Biden signed the Executive order on "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence."[16] Noting the importance of maximizing the benefits of open foundation models while managing and mitigating the attendant risks, section 4.6 the

Executive order tasked the Secretary of Commerce, acting through NTIA and in consultation with the Secretary of State, with soliciting feedback "from the private sector, academia, civil society, and other stakeholders through a public consultation process on the potential risks, benefits, other implications, and appropriate policy and regulatory approaches related to dual-use foundation models for which the model weights are widely available."[17] As required by the Executive order, the Secretary of Commerce, through NTIA, and in consultation with the Secretary of State, will author a report to the President on the "potential benefits, risks, and implications of dual-use foundation models for which the model weights are widely available, as well as policy and regulatory recommendations pertaining to those models."[18]

In particular, the Executive order asks NTIA to consider risks and benefits of dual-use foundation models with weights that are "widely available."[19] Likewise, "openness" or "wide availability" of model weights are also terms without clear definition or consensus. There are gradients of "openness," ranging from fully "closed" to fully "open."[20] There is also more information needed to detail the relationship between openness and the wide availability of both model weights and open foundation models more generally. This could include, for example, information about what types of licenses and distribution methods are available or could be available for open foundation models, and how such licenses and distribution methods fit within an understanding of openness and wide availability.[21]

NTIA also requests input on any potential regulatory models, either voluntary or mandatory, that could maintain and potentially increase the benefits and/or mitigate the risks of dual use foundation models with widely available model weights. We seek input as to different kinds of regulatory structures that could deal with not only the large scale of these foundation models, but also the declining level of

computing resources needed to fine-tune and retrain them.

Definitions

This Request for Comment uses the terms defined in sec. 3 of the Executive order. In addition, we use broader terms interchangeably for both ease of understanding and clarity, as set forth below. "Artificial intelligence" or "AI" refer to a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions, influencing real or virtual environments.[22] Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments, abstract such perceptions into models through analysis in an automated manner, and use model inference to formulate options for information or action.

Foundation models are typically defined as, "powerful models that can be fine-tuned and used for multiple purposes."[23] Under the Executive order, a "dual-use foundation model" is "an AI model that is trained on broad data; generally uses self-supervision; contains at least tens of billions of parameters; is applicable across a wide range of contexts; and that exhibits, or could be easily modified to exhibit, high levels of performance at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matters . . . ."[24] Both definitions of "foundation model" and of "dual-use foundation model"—highlight the key trait of these models, that they can be used in a number of ways.[25]

"Generative AI can be understood as a form of AI model specifically intended to produce new digital material as an output (including text, images, audio, video, software code), including when such AI models are used in applications and their user interfaces."[26] The term "generative AI" refers to a class of AI models built on foundation models

[11] Id.

[12] Id.

[13] For example, researchers have found ways to get both black box large language models as well as more open models to produce objectionable content through adversarial attacks. *See e.g.,* Andy Zou et al., Universal and Transferable Adversarial Attacks on Aligned Language Models, arXiv:2307.15043 (July 27, 2023). *https://arxiv.org/abs/2307.15043* ("Surprisingly, we find that the adversarial prompts generated by our approach are quite transferable, including to black-box, publicly released LLMs . . . When doing so, the resulting attack suffix is able to induce objectionable content in the public interfaces to ChatGPT, Bard, and Claude, as well as open source LLMs such as LLaMA–2–Chat, Pythia, Falcon, and others.").

[14] See *e.g.,* Zoe Brammer, How Does Access Impact Risk? Assessing AI Foundation Model Risk Along a Gradient of Access, The Institute for Security and Technology (December 2023) *https:// securityandtechnology.org/wp-content/uploads/ 2023/12/How-Does-Access-Impact-Risk-Assessing- AI-Foundation-Model-Risk-Along-A-Gradient-of- Access-Dec-2023.pdf.*

[15] Id and *see e.g.* Pranshu Verma, The rise of AI fake news is creating a 'misinformation superspreader', Washington Post (December 17, 2023) *https://www.washingtonpost.com/ technology/2023/12/17/ai-fake-news- misinformation/.*

[16] E.O. 14110, 88 FR 75191 (November 1, 2023).

[17] Id.

[18] Id.

[19] E.O. 14110, 88 FR 75191 (November 1, 2023).

[20] *See, e.g.,* Irene Solaiman, The Gradient of Generative AI Release: Methods and Considerations, arXiv:2302.04844v1 (February 5, 2023) *https://arxiv.org/pdf/2302.04844.pdf;* Bommasani et al., *supra* note 9.

[21] *See, e.g.,* Carlos Munoz Ferrandis, OpenRAIL: Towards open and responsible AI licensing frameworks, Hugging Face Blog (August 31, 2022) *https://huggingface.co/blog/open_rail;* Danish Contractor et al., Behavioral Use Licensing for Responsible AI, arXiv:2011.03116v2 (October 20, 2022) *https://arxiv.org/pdf/2011.03116.pdf.*

[22] E.O. 14110, 88 FR 75191 (November 1, 2023).

[23] *See, e.g.,* "A foundation model is any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (*e.g.,* fine-tuned) to a wide range of downstream tasks[.]" Rishi Bommasani et al., On the Opportunities and Risks of Foundation Models, arXiv:2108.07258v3 (July 12, 2022). *https://arxiv.org/pdf/ 2108.07258.pdf.*

[24] E.O. 14110, 88 FR 75191 (November 1, 2023).

[25] Id.

[26] G7 Hiroshima Process on Generative Artificial Intelligence (AI) Towards a G7 Common Understanding on Generative AI, Organisation for Economic Co-operation and Development (OECD) (September 7, 2023) *https://www.oecd-ilibrary.org/ docserver/bf3c0c60-en.pdf?expires=1705032283& id=id&accname=guest&checksum=85A1D78C60AC 6D8BBFBF2514CB7F2A5D.*

''that emulate the structure and characteristics of input data in order to generate derived synthetic content.'' [27] Chatbots like ChatGPT, large language models like BLOOM, and image generators like Midjourney are all examples of generative AI.

This Request for Comment is particularly focused on the wide availability, such as being publicly posted online, of foundation model weights. ''Model weights'' are ''numerical parameter[s] within an AI model that help [. . .] determine the model's output in response to inputs.'' [28] In addition to model weights, there are other ''components'' of an AI model, including training data, code, or other elements, which are involved in its development or use, and may or may not be made widely available.

The Executive order directs NTIA to focus on dual-use foundation models that were trained on broad data; generally use self-supervision; contain at least tens of billions of parameters; are applicable across a wide range of contexts; and exhibit, or could be easily modified to exhibit, high levels of performance at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matter.[29] NTIA also remains interested in the discussion of models that fall outside of the scope of this Request for Comments in order to better understand the current landscape and potential impact of regulatory or policy actions.

Instructions for Commenters

Through this Request for Comment, we hope to gather information on the following questions. These are not exhaustive, and commenters are invited to provide input on relevant questions not asked below. Commenters are not required to respond to all questions. When responding to one or more of the questions below, please note in the text of your response the number of the question to which you are responding. Commenters should include a page number on each page of their submissions. Commenters are welcome to provide specific actionable proposals, rationales, and relevant facts.

Please do not include in your comments information of a confidential nature, such as sensitive personal information or proprietary information. All comments received are a part of the public record and will generally be posted to *Regulations.gov* without change. All personal identifying

information (*e.g.,* name, address) voluntarily submitted by the commenter may be publicly accessible.

Questions

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

a. Is there evidence or historical examples suggesting that weights of models similar to currently-closed AI systems will, or will not, likely become widely available? If so, what are they?

b. Is it possible to generally estimate the timeframe between the deployment of a closed model and the deployment of an open foundation model of similar performance on relevant tasks? How do you expect that timeframe to change? Based on what variables? How do you expect those variables to change in the coming months and years?

c. Should ''wide availability'' of model weights be defined by level of distribution? If so, at what level of distribution (*e.g.,* 10,000 entities; 1 million entities; open publication; etc.) should model weights be presumed to be ''widely available''? If not, how should NTIA define ''wide availability?''

d. Do certain forms of access to an open foundation model (web applications, Application Programming Interfaces (API), local hosting, edge deployment) provide more or less benefit or more or less risk than others? Are these risks dependent on other details of the system or application enabling access?

i. Are there promising *prospective* forms or modes of access that could strike a more favorable benefit-risk balance? If so, what are they?

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

a. What, if any, are the risks associated with widely available model weights? How do these risks change, if at all, when the training data or source code associated with fine tuning, pretraining, or deploying a model is simultaneously widely available?

b. Could open foundation models reduce equity in rights and safety-impacting AI systems (*e.g.,* healthcare, education, criminal justice, housing, online platforms, etc.)?

c. What, if any, risks related to privacy could result from the wide availability of model weights?

d. Are there novel ways that state or non-state actors could use widely available model weights to create or exacerbate security risks, including but not limited to threats to infrastructure,

public health, human and civil rights, democracy, defense, and the economy?

i. How do these risks compare to those associated with closed models?

ii. How do these risks compare to those associated with other types of software systems and information resources?

e. What, if any, risks could result from differences in access to widely available models across different jurisdictions?

f. Which are the most severe, and which the most likely risks described in answering the questions above? How do these set of risks relate to each other, if at all?

3. What are the benefits of foundation models with model weights that are widely available as compared to fully closed models?

a. What benefits do open model weights offer for competition and innovation, both in the AI marketplace and in other areas of the economy? In what ways can open dual-use foundation models enable or enhance scientific research, as well as education/training in computer science and related fields?

b. How can making model weights widely available improve the safety, security, and trustworthiness of AI and the robustness of public preparedness against potential AI risks?

c. Could open model weights, and in particular the ability to retrain models, help advance equity in rights and safety-impacting AI systems (*e.g.,* healthcare, education, criminal justice, housing, online platforms etc.)?

d. How can the diffusion of AI models with widely available weights support the United States' national security interests? How could it interfere with, or further the enjoyment and protection of human rights within and outside of the United States?

e. How do these benefits change, if at all, when the training data or the associated source code of the model is simultaneously widely available?

4. Are there other relevant components of open foundation models that, if simultaneously widely available, would change the risks or benefits presented by widely available model weights? If so, please list them and explain their impact.

5. What are the safety-related or broader technical issues involved in managing risks and amplifying benefits of dual-use foundation models with widely available model weights?

a. What model evaluations, if any, can help determine the risks or benefits associated with making weights of a foundation model widely available?

b. Are there effective ways to create safeguards around foundation models,

[27] E.O. 14110, 88 FR 75191 (November 1, 2023).
[28] Id.
[29] Id.

either to ensure that model weights do not become available, or to protect system integrity or human well-being (including privacy) and reduce security risks in those cases where weights are widely available?

c. What are the prospects for developing effective safeguards in the future?

d. Are there ways to regain control over and/or restrict access to and/or limit use of weights of an open foundation model that, either inadvertently or purposely, have already become widely available? What are the approximate costs of these methods today? How reliable are they?

e. What if any secure storage techniques or practices could be considered necessary to prevent unintentional distribution of model weights?

f. Which components of a foundation model need to be available, and to whom, in order to analyze, evaluate, certify, or red-team the model? To the extent possible, please identify specific evaluations or types of evaluations and the component(s) that need to be available for each.

g. Are there means by which to test or verify model weights? What methodology or methodologies exist to audit model weights and/or foundation models?

6. What are the legal or business issues or effects related to open foundation models?

a. In which ways is open-source software policy analogous (or not) to the availability of model weights? Are there lessons we can learn from the history and ecosystem of open-source software, open data, and other ''open'' initiatives for open foundation models, particularly the availability of model weights?

b. How, if at all, does the wide availability of model weights change the competition dynamics in the broader economy, specifically looking at industries such as but not limited to healthcare, marketing, and education?

c. How, if at all, do intellectual property-related issues—such as the license terms under which foundation model weights are made publicly available—influence competition, benefits, and risks? Which licenses are most prominent in the context of making model weights widely available? What are the tradeoffs associated with each of these licenses?

d. Are there concerns about potential barriers to interoperability stemming from different incompatible ''open'' licenses, *e.g.,* licenses with conflicting requirements, applied to AI components? Would standardizing

license terms specifically for foundation model weights be beneficial? Are there particular examples in existence that could be useful?

7. 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? What kind of entities should take a leadership role across which features of governance?

a. What security, legal, or other measures can reasonably be employed to reliably prevent wide availability of access to a foundation model's weights, or limit their end use?

b. How might the wide availability of open foundation model weights facilitate, or else frustrate, government action in AI regulation?

c. When, if ever, should entities deploying AI disclose to users or the general public that they are using open foundation models either with or without widely available weights?

d. What role, if any, should the U.S. government take in setting metrics for risk, creating standards for best practices, and/or supporting or restricting the availability of foundation model weights?

i. Should other government or non-government bodies, currently existing or not, support the government in this role? Should this vary by sector?

e. What should the role of model hosting services (*e.g.,* HuggingFace, GitHub, etc.) be in making dual-use models with open weights more or less available? Should hosting services host models that do not meet certain safety standards? By whom should those standards be prescribed?

f. Should there be different standards for government as opposed to private industry when it comes to sharing model weights of open foundation models or contracting with companies who use them?

g. What should the U.S. prioritize in working with other countries on this topic, and which countries are most important to work with?

h. What insights from other countries or other societal systems are most useful to consider?

i. Are there effective mechanisms or procedures that can be used by the government or companies to make decisions regarding an appropriate degree of availability of model weights in a dual-use foundation model or the dual-use foundation model ecosystem? Are there methods for making effective decisions about open AI deployment that balance both benefits and risks? This may include responsible capability

scaling policies, preparedness frameworks, et cetera.

j. Are there particular individuals/ entities who should or should not have access to open-weight foundation models? If so, why and under what circumstances?

8. 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?

a. How should these potentially competing interests of innovation, competition, and security be addressed or balanced?

b. Noting that E.O. 14110 grants the Secretary of Commerce the capacity to adapt the threshold, is the amount of computational resources required to build a model, such as the cutoff of 10^{26} integer or floating-point operations used in the Executive order, a useful metric for thresholds to mitigate risk in the long-term, particularly for risks associated with wide availability of model weights?

c. Are there more robust risk metrics for foundation models with widely available weights that will stand the test of time? Should we look at models that fall outside of the dual-use foundation model definition?

9. What other issues, topics, or adjacent technological advancements should we consider when analyzing risks and benefits of dual-use foundation models with widely available model weights?

Dated: February 20, 2024.

Stephanie Weiner,

Chief Counsel, National Telecommunications and Information Administration.

[FR Doc. 2024–03763 Filed 2–23–24; 8:45 am]

BILLING CODE 3510–60–P

DEPARTMENT OF DEFENSE

Defense Acquisition Regulations System

[Docket No. 2024–0006; OMB Control No. 0750–0004]

Information Collection Requirement; Defense Federal Acquisition Regulation Supplement; Assessing Contractor Implementation of Cybersecurity Requirements

AGENCY: Defense Acquisition Regulations System; Department of Defense (DOD).

ACTION: Notice and request for comments regarding a proposed