epic.org

Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, DC 20036, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

COMMENTS OF THE ELECTRONIC PRIVACY INFORMATION CENTER

to the

DEPARTMENT OF COMMERCE BUREAU OF INDUSTRY AND SECURITY

Request for Comment: Establishment of Reporting Requirements for the Development of Advanced Artificial Intelligence Models and Computing Clusters

No. 240905-0231; RIN 0694-AJ55

October 11, 2024

The Electronic Privacy Information Center (EPIC) submits these comments in response to the Bureau of Industry and Security at the Department of Commerce (Bureau)'s Request for Comment on the proposed rule amending the Bureau's Industrial Base Surveys – Data Collections regulations by establishing reporting requirements for the development of advanced artificial intelligence (AI) models and computing clusters.[1]

EPIC is a public interest research center based in Washington, D.C., established in 1994 to secure the fundamental right to privacy in the digital age for all people through advocacy, research, and litigation.[2] We work to center human rights and dignity in AI policy and development, ensuring new technologies are subject to democratic governance and accountability.[3] Over the last decade, EPIC has consistently advocated for the adoption of clear, commonsense, actionable AI regulations, across the country.[4] EPIC has published extensive research on the risks and harms of emerging AI

---

[1] 89 Fed. Reg. 73612 (Sept. 11, 2024).
[2] *About Us,* EPIC (2023), https://epic.org/about/.
[3] *See, e.g., AI and Human Rights,* EPIC (2023), https://epic.org/issues/ai/; *AI and Human Rights: Criminal Legal System,* EPIC (2023), https://epic.org/issues/ai/ai-in-the-criminal-justice-system/; "Outsourced & Automated: How AI Companies Have Taken Over Government Decision-Making," EPIC (2023), available at https://epic.org/outsourced-automated; Letter from EPIC to President Biden and Vice President Harris on Ensuring Adequate Federal Workforce and Resources for Effective AI Oversight (Oct. 24, 2023), available at https://epic.org/wp-content/uploads/2023/10/EPIC-letter-to-White-House-re-AI-workforce-and-resources-Oct-2023.pdf; Comments on Draft Documents Responsive to NIST's Assignments Under Executive order 14110 (Sections 4.1, 4.5, and 11), EPIC (June 2, 2024), available at https://epic.org/epic-pushes-nist-to-focus-its-approach-to-generative-ai-risks-around-who-and-how-ai-harms/; Comments on the NIST Artificial Intelligence Risk Management Framework: Second Draft, EPIC (Sept. 28, 2022), available at https://epic.org/wp-content/uploads/2022/09/EPIC-Comments-NIST-RMF-09-28-22.pdf.
[4] *See, e.g.,* Press Release, EPIC, EPIC Urges DC Council to Pass Algorithmic Discrimination Bill (Sept. 23,

technologies, such as generative AI,[5] as well as how government agencies develop, procure, and use AI systems.[6] Finally, EPIC has repeatedly drawn attention to misuse, privacy, and human rights risks inherent in dual-use foundation models specifically.[7]

The Bureau's proposed rule would be one of the first established requirements on reporting for the AI industry. This rule sets the tone for multiple other regulatory bodies and agencies in how to approach reporting, transparency, assessment, and enforcement for AI. EPIC applauds the Bureau for taking a proactive reporting approach in this instance. The AI industry has lacked transparency about its development and deployment of models and other AI systems, meaning evaluation of efficacy and risk can only meaningfully occur after the models and systems are already in use and actively generating harms.

We agree with the Bureau that it is essential to ensure AI models used by the U.S. Government must operate in a "safe and reliable manner," particularly due to the severe consequences of failures of these systems, including potential loss of life.[8] Risks of unfitness, inconsistency, vulnerabilities, and exploitation in these systems can lead to catastrophic harms that cannot be remedied. However, there are significant sociotechnical risks, particularly of generative AI, that cannot be uncoupled from these national security risks and so must be considered when forming the Bureau's proposed rule.

Many foundation models are built using the largest datasets possible. This often involves building training datasets through indiscriminate web scraping and purchasing data from data

---

2022), https://epic.org/epic-urges-dc-council-to-pass-algorithmic-discrimination-bill/; EPIC, Comments to the Patent and Trademark Office on Intellectual Property Protection for Artificial Intelligence Innovation (Jan. 10, 2020), https://epic.org/wp-content/uploads/apa/comments/EPIC-USPTO-Jan2020.pdf; EPIC, Comments on the Department of Housing and Urban Development's Implementation of the Fair Housing Act's Disparate Impact Standard (Oct. 18, 2019), https://epic.org/wp-content/uploads/apa/comments/EPIC-HUD-Oct2019.pdf.

[5] "Generating Harms: Generative AI's Impact & Paths Forward," EPIC (2023) and "Generating Harms II: Generative AI's New & Continued Impacts," EPIC (2024), both available at https://epic.org/generating-harms/.

[6] Outsourced & Automated Report; "Screened & Scored in the District of Columbia," EPIC (2022), available at https://epic.org/wp-content/uploads/2022/11/EPIC-Screened-in-DC-Report.pdf.

[7] *See, e.g.,* Comments on U.S. Artificial Intelligence Safety Institute's Draft Document: Managing Misuse Risk for Dual-Use Foundation Models, EPIC (Sept. 9, 2024), available at https://epic.org/documents/epic-comments-to-nist-on-managing-the-risks-of-misuse-with-ai-foundation-models/; Comments on Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights, EPIC (March 27, 2024), available at https://epic.org/epic-urges-the-ntia-to-tackle-privacy-harms-bias-and-regulatory-hurdles-in-new-comment-on-ai-model-openness/.

[8] 89 Fed. Reg. 73612 (Sept. 11, 2024).

brokers.[9] The volume of data used in these cases often makes it impossible to fully curate or clean when it comes to accuracy, bias, or personal information. Once a model is trained on data, it cannot be easily removed or unlearned.[10] In addition, these massive datasets, which often contain personal and sensitive information, are vulnerable to attack – we have already seen examples of effective membership inference attacks and attribute inference attacks.[11] Security concerns in these model extend beyond the training datasets to the structure of the models themselves, implementation, access, modification options, and more.

While foundation models are subject to many risks, we will highlight just two of the most potentially damaging here: bias and accuracy. Training data often reflects human and historical biases.[12] These biases have already produced tangible harms across multiple industries and uses, including loss of employment opportunities, biased medical treatments, and racial and gender bias in facial recognition systems.[13] While some techniques have been proposed to debias data and models,[14]

---

[9] *See* Müge Fazlioglu, *Training AI on Personal Data Scraped from the Web*, IAPP (Nov. 8, 2023), https://iapp.org/news/a/training-ai-on-personal-data-scraped-from-the-web/; Thomas Claburn, *How to Spot OpenAI's Crawler Bot and Stop it Slurping Sites for Training Data*, Register (Aug. 8, 2023), https://www.theregister.com/2023/08/08/openai_scraping_software/; Sara Morrison, *The Tricky Truth About How Generative AI Uses Your Data,* Vox (July 27, 2023), https://www.vox.com/technology/2023/7/27/23808499/ai-openai-google-meta-data-privacy-nope; Evan Weinberger, *Data Brokers Eyed by CFPB for Selling Sensitive Info for Ads, AI,* Bloomberg Law (Aug. 15, 2023), https://news.bloomberglaw.com/banking-law/data-brokers-eyed-by-cfpb-for-selling-sensitive-info-for-ads-ai.

[10] *See* Liwei Song & Prateek Mittal, *Systematic Evaluation of Privacy Risks of Machine Learning Models*, 30 Proc. USENIX Sec. Symp. 2615, 2615 (2021). Hurdles to unlearning data are at the core of recent FTC cases requiring AI model deletion. *See* Jevan Hutson & Ben Winters, *America's Next 'Stop Model!': Model Deletion*, 8 Geo. L. Tech. Rev. 125, 128–134 (2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4225003.

[11] *See* Song & Mittal, *supra* note 10 at 2629; Reza Shokri et al., *Membership Inference Attacks Against Machine Learning Models,* 2017 IEEE Sump. On Sec. & Priv. 3, https://ieeexplore.ieee.org/document/7958568; Bargav Jayaraman & David Evans, *Are Attribute Inference Attacks Just Imputation?,* arXiv (Sept. 2, 2022), https://arxiv.org/pdf/2209.01292.pdf.

[12] *See* IBM Data and AI Team, *Shedding Light on AI Bias with Real World Examples*, IBM (Oct. 16, 2023), https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/; Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 Proc. Mach. Learning Rsch. 1–15 (2018), https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

[13] *See, e.g.*, Leon Yin et al., *OpenAI's GPT is a Recruiter's Dream Tool. Tests Show There's Racial Bias*, Bloomberg (March 7, 2024); Jesutofunmi A. Omiye et al., *Large Language Models Propagate Race-Based Medicine*, 6 npj Digit. Med. 1–3 (2023), https://www.nature.com/articles/s41746-023-00939-z; Drew Harwell, *Federal Study Confirms Racial Bias of Many Facial-Recognition Systems, Casts Doubt on Their Expanding Use*, Wash. Post (Dec. 19, 2019), https://www.washingtonpost.com/technology/2019/12/19/federal-study-confirms-racial-bias-many-facial-recognition-systems-casts-doubt-their-expanding-use/.

[14] *See* Hyojin Bahng et al., *Learning De-biased Representations with Biased Representations*, 37 Proc. Int'l Conf. on Mach. Learning 1 (2020), https://proceedings.mlr.press/v119/bahng20a/bahng20a.pdf; Robert

all of these depend on developers' ability to accurately identify the source of bias when training an AI model. This high level risk and the challenges to mitigate it should be taken into account in reporting requirements for models. Several governmental bodies, including NIST, the Government Accountability Office, OSTP, and others, have published reports on key issues with AI, including bias and discrimination.[15]

Foundation models confront substantial challenges regarding accuracy that require vigilant human intervention to counteract.[16] Despite rapid advancements in AI language generation, image recognition, and decision making, AI systems are intrinsically susceptible to fallibility, particularly in intricate or nuanced scenarios.[17] This is particularly dangerous in government application and interaction, as the consequences of inaccuracy could be catastrophic practically, politically, and individually.

For the above reasons and many more identified in the supplemental resources included below, we recommend that the Bureau include in its reporting requirements, at minimum, (i) complete information about the origin of data included in training datasets; (ii) the data review and curation process for this training data, including any efforts at accuracy and bias checks; (iii) regular risk assessments for inputs, internal functions, and outputs of foundation models; (iv) all internal and

---

Geirhos et al., *ImageNet-Trained CNNs are Biased Towards Texture; Increasing Shape Bias Improves Accuracy and Robustness,* 2019 Proc. Int'l Conf. on Learning Representations 7, 17, https://openreview.net/pdf?id=Bygh9j09KX; Haohan Wang et al., *Learning Robust Representations by Projecting Superficial Statistics Out*, 2019 Proc. Int'l Conf. on learning Representations 8–9, https://openreview.net/pdf?id=rJEjjoR9K7; Faisal Kamiran & Toon Calders, *Data Preprocessing Techniques for Classification Without Discrimination*, 33 Knowledge Info. Sys. 1, 2, 16–18 (2012), https://link.springer.com/article/10.1007/s10115-011-0463-8.

[15] *See generally* GAO, GAO-21-519SP, Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities (2021), https://www.gao.gov/assets/gao-21-519sp.pdf; GAO, GAO-21-7SP, Artificial Intelligence in Health Care: Benefits and Challenges of Technologies to Augment Patient Care (2020), https://www.gao.gov/products/gao-21-7sp; Blueprint for an AI Bill of Rights at 23–29; Reva Schwartz et al., NIST, Towards a Standard for Identifying and Managing Bias in Artificial Intelligence (NIST Special Publ'n 1270, March 15, 2022), https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf.

[16] Arun Shastri, *Generative AI Errs Differently Than Classical AI*, Forbes (Sept. 4, 2023), https://www.forbes.com/sites/arunshastri/2023/09/04/generative-ai-errs-differently-than-classical-ai/.

[17] Oceane Duboust, *Unreliable Research Assistant? False Outputs from AI Chatbots Pose Risk to Science, Report Says*, Euronews (Nov. 20, 2023), https://www.euronews.com/next/2023/11/20/unreliable-research-assistant-false-outputs-from-ai-chatbots-pose-risk-to-science-report-s; Tate Ryan-Mosley, *Catching Bad Content in the Age of AI*, MIT Tech. Rev. (May 15, 2023), https://www.technologyreview.com/2023/05/15/1073019/catching-bad-content-in-the-age-of-ai/ (explaining how generative AI's difficulty handling nuanced information presents problems for content moderation).

external audit documentation; (v) policies, plans, and other documentation related to risk mitigation; (vi) use case inventories; and (vii) regular tests for fitness, functionality, and reliability.

We appreciate the opportunity to comment on the Bureau's proposed rule and are available for any further engagement regarding these or other issues raised within our comment or the supplemental resources provided. We hope that the proposed rule ensures meaningful oversight and transparency for the development of advanced AI models and computing clusters and leads to safe, equitable, and trustworthy government AI use.

Respectfully submitted,

/s/ Calli Schroeder
Calli Schroeder
Senior Counsel and AI & Human Rights Lead

ELECTRONIC PRIVACY
INFORMATION CENTER (EPIC)
1519 New Hampshire Ave. NW
Washington, D.C. 20036
202-483-1140

ADDITIONAL RESOURCES

REPORTS

"Outsourced & Automated: How AI Companies Have Taken Over Government Decision-Making," EPIC (2023), available at https://epic.org/outsourced-automated

"Generating Harms: Generative AI's Impact & Paths Forward," EPIC (2023) and "Generating Harms II: Generative AI's New & Continued Impacts," EPIC (2024), both available at https://epic.org/generating-harms/

"Screened & Scored in the District of Columbia," EPIC (2022), available at https://epic.org/wp-content/uploads/2022/11/EPIC-Screened-in-DC-Report.pdf

COMMENTS

Comments on the U.S. Artificial Intelligence Safety Institute's Draft Document: Managing Misuse Risk for Dual-Use Foundation Models, EPIC (September 9, 2024), available at https://epic.org/documents/epic-comments-to-nist-on-managing-the-risks-of-misuse-with-ai-foundation-models/

Comments to the Privacy and Civil Liberties Oversight Board on The Role of Artificial Intelligence in Counterterrorism, EPIC (July 1, 2024), available at https://epic.org/documents/comments-of-epic-to-pclob-ai-in-counterterrorism-july-2024/

Comments on Draft Documents Responsive to NIST's Assignments Under Executive order 14110 (Sections 4.1, 4.5, and 11), EPIC (June 2, 2024), available at https://epic.org/epic-pushes-nist-to-focus-its-approach-to-generative-ai-risks-around-who-and-how-ai-harms/

Comments to NTIA on Dual Use Foundation Artificial Intelligence Models with Widely Available Model Weights, EPIC (March 27, 2024), available at https://epic.org/epic-urges-the-ntia-to-tackle-privacy-harms-bias-and-regulatory-hurdles-in-new-comment-on-ai-model-openness/

Comments to the Office of Management and Budget on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence Draft Memorandum, EPIC (Dec. 5, 2023), available at https://epic.org/wp-content/uploads/2023/12/EPIC-OMB-AI-Guidance-Comments-120523-1.pdf

Comments on the NIST Artificial Intelligence Risk Management Framework: Second Draft, EPIC (September 28, 2022), available at https://epic.org/wp-content/uploads/2022/09/EPIC-Comments-NIST-RMF-09-28-22.pdf