Submitted by Susan Chertkof on behalf of Recording Industry Association of America, Inc. ("RIAA")

Before the
DEPARTMENT OF COMMERCE
Bureau of Industry and Security
Washington, D.C.

In the Matter of: Establishment of Reporting Requirements for the Development of Advanced Artificial Intelligence Models And Computing Clusters	Docket No. 240905-0231 RIN 0694-AJ55 Submitted October 11, 2024

COMMENTS OF THE
RECORDING INDUSTRY ASSOCIATION OF AMERICA, INC.

I. INTRODUCTION

The Recording Industry Association of America, Inc. ("RIAA") is pleased to provide comments to the Department of Commerce ("Department"), Bureau of Industry and Security ("BIS"), in response to the proposed rule ("NPRM") regarding the Establishment of Reporting Requirements for the Development of Advanced Artificial Intelligence Models and Computing Clusters published at 89 Fed. Reg. 73,612 (Sept. 11, 2024).

The Recording Industry Association of America is the trade organization that supports and promotes the creative and commercial vitality of music labels in the United States, the most vibrant recorded music community in the world. RIAA's membership – which includes several hundred companies, ranging from small-to-medium-sized enterprises to global businesses – creates, manufactures, and/or distributes sound recordings representing the majority of all lawfully recorded music consumption in the United States. In support of its mission, the RIAA works to protect the intellectual property and First Amendment rights of artists and music labels; conducts consumer, industry, and technical research; and monitors and reviews state and federal laws, regulations, and policies.

On behalf of our members, we have been closely following developments in the artificial intelligence ("AI") space and have filed comments in that regard with a number of federal agencies. In each instance, we have consistently encouraged the U.S government and AI companies to establish responsible recordkeeping and reporting policies to ensure the development and deployment of AI, including dual-use foundation models, is done in a

safe, trustworthy, and responsible manner that complies with all applicable laws, including intellectual property laws.[1] With this goal in mind, we offer the following comments on the proposed rule.

II. Comments

The proposed rule would amend the BIS's Industrial Base Surveys—Data Collections regulations by establishing reporting requirements for the development of advanced AI models and computing clusters under Executive Order 14110 of October 30, 2023, "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" (the "Executive Order").[2] As explained in the NPRM, in order for the government to "ensure [that] dual-use foundation models produced by U.S. companies are available to the defense industrial base,"[3] the government needs information about "how many U.S. companies are developing, have plans to develop, or have the computing hardware necessary to develop dual-use foundation models, as well as information about the characteristics of dual-use foundation models under development."[4] The government also needs to ensure that any dual-use foundation models adopted by the defense industrial base will "operate in a safe and reliable manner."[5]

To address these concerns, the proposed rule would create a notification and reporting process for "companies developing or intending to develop dual-use foundation AI models and for companies, individuals or other organizations or entities that acquire, develop, or possess computing clusters that meet technical conditions issued by the Department."[6] All persons covered by the proposed rule would be required to identify themselves and report certain information to the BIS on a quarterly basis.[7]

As BIS considers exactly what information to collect concerning dual-use foundation models, it is vital that relevant stakeholders be required to regularly provide detailed data identifying exactly what materials were used to train or fine tune the models, and the provenance of those training materials.

As the Department is likely aware, some generative AI companies have "scraped" the internet to collect training materials, without regard to a source site's terms of service or legality, and without authorization from the owners of the content being copied. Such

---

[1] See, e.g., the following comments we (together with others) have submitted to U.S. agencies: National Telecommunications and Information Administration ("NTIA"), Docket No. NTIA-2023-0005, https://www.regulations.gov/comment/NTIA-2023-0005-1277 ("NTIA Comments") and Office of Science and Technology Policy ("OSTP"), Docket No. OSTP-TECH-2023-0007, https://www.regulations.gov/comment/OSTP-TECH-2023-0007-0231 ("OSTP Comments").

[2] 2023-24283.pdf (govinfo.gov)

[3] NPRM at 73613.

[4] Id.

[5] Id.

[6] NPRM at 73614.

[7] Id.

scraping raises concerns about the overall safety and reliability of products and services that are built using the scraped inputs. The training material so taken may not be from an authorized source, may not be accurate, or may have been intentionally modified, and the use of such materials can introduce risks into the model and its output. These include intellectual property risks, privacy risks, cyber vulnerabilities, fraud, safety, and other risks.

To minimize these risks, the regulations that are the subject of this rulemaking should mandate that the relevant stakeholders, including developers and deployers of dual-use foundation models and those that develop or aggregate datasets for use for such foundation models, maintain complete and detailed documentation about: (1) what specific materials were ingested to develop a dual-use foundation model (or to fine-tune a pre-trained dual-use foundation model) and in what manner; (2) the source and provenance of such materials, including whether any licenses or authorizations were sought or obtained to authorize such use; (3) the articulated rationale for selecting and using the materials ingested for the dual-use foundation model's development; (4) the individual or organization responsible for the dual-use foundation model (including who is responsible for ingesting the materials, who is responsible for any fine-tuning of the model, who is deploying the model, etc.); and (5) the extent to which each copyrighted work in the training materials was relied upon in order to generate specific outputs from the AI model.[8] For transparency, the relevant stakeholders should be required to disclose this and other relevant information to BIS.

Both NTIA recommendations and NIST best practices agree with this approach. As NTIA recommends, the information "should include documentation about AI system models, architecture, data, performance, limitations, appropriate use, and testing,"[9] noting that "[g]reater transparency about, for example, AI system models, architecture, training data, input and output data, performance, limitations, appropriate use, and testing should be provided to relevant audiences."[10] Similarly, NIST recommends that the relevant stakeholders establish "transparency policies and processes for documenting the origin and history of training data and generated data for GAI applications to advance digital content transparency, while balancing the proprietary nature of training approaches."[11] In developing its regulations, BIS should ensure that all relevant stakeholders engage in appropriate, detailed recordkeeping and transparency consistent with NTIA's recommendations and NIST's best practices.

---

[8] As noted by the National Institute of Standards and Technology (NIST), "[m]aintaining the provenance of training data and supporting attribution of the AI system's decisions to subsets of training data can assist with both transparency and accountability."Nat'l Inst. of Standards and Tech., Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1, 16 (2023), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[9] National Telecommunications and Information Administration, *NTIA Artificial Intelligence Accountability Report*, March 2024, p. 2, https://www.ntia.gov/sites/default/files/publications/ntia-ai-report-final.pdf.

[10] Id. at p. 3.

[11] Id., Action ID GV-1.2-001, p. 14.

In connection with this, BIS should examine and report on any unauthorized copyrighted content copied into the training datasets used by relevant stakeholders.[12] This is an important moment for the US government to push the nascent AI industry to respect the IP belonging to others, just as they desire their own IP to be protected.

Thank you for the opportunity to comment on this important issue with critical ramifications for the integrity of our defense industrial base. If you have additional questions about our views, we would be happy to answer them.

Respectfully Submitted,

Susan B. Chertkof
Senior Vice President, Legal & Regulatory Affairs
Recording Industry Association of America, Inc.
1000 F Street, N.W.
Floor 2
Washington, D.C. 20004
(v) 202-775-0101
(f) 202-775-7253
schertkof@riaa.com

[12] Of course, BIS should defer on any discussion of AI and fair use to other bodies tasked with undertaking such an analysis.