BEFORE THE
DEPARTMENT OF ENERGY

Request for Information; **Partnerships for Transformational** **Artificial Intelligence Models**

RFI Number: DE-ASCR-26-0001

The Copyright Alliance appreciates the opportunity to submit the following comments in response to the request for information (RFI) published by the Department of Energy (DOE) in the Federal Register on December 5, 2025, requesting input on "how DOE should best structure and enable partnerships to curate DOE scientific data across the National Laboratory complex for use in artificial intelligence (AI) models" and on "using this data to develop self-improving AI models for science and engineering to advance scientific discovery, energy, and national security."

The Copyright Alliance is a non-profit, non-partisan public interest and educational organization representing the copyright interests of over 2 million individual creators and over 15,000 organizations in the United States, across the spectrum of copyright disciplines. The Copyright Alliance is dedicated to advocating policies that promote and preserve the value of copyright, and to protecting the rights of creators and innovators. The individual creators and organizations that we represent rely on copyright law to protect their creativity, efforts, and investments in the creation and distribution of new copyrighted works for the public to enjoy.

The Copyright Alliance and its members support the establishment of a public-private consortium to curate the scientific data of DOE across the National Laboratory complex for use in AI models so long as such curation and resultant uses are based on a comprehensive intellectual property (IP)

framework and respect the tenets and provisions in U.S. copyright law.[1] Consistent with these views, we submit these comments to urge that any new policies aimed at structuring and enabling partnerships to curate DOE scientific data for use in AI models (i) be developed and implemented with a respect for and in recognition of longstanding copyright laws and policies that make America the global leader in scientific discovery and technological progress, and (ii) do not disrupt the critical scientific and research publishing industries, the incentives upon which they are based, or the licensing markets that sustain them.

The U.S IP system, enshrined in the Constitution, incentivizes the dissemination of new innovations and publications, but in a way that protects their value—often via licensing contracts that create a relationship between parties and ensure that AI systems are being trained on high-quality and reliable content. It is therefore essential that mobilization of the DOE, the National Laboratories, and private partners in executing America's AI Action Plan[2] and the Genesis Mission[3] be done with a respect for copyright law and the exclusive rights that incentivize the creation and dissemination of copyrighted works. This includes avoiding broad definitions of "data" that would allow for unbridled government use or licensing (including for AI training) of copyright protected works created by private sector entities.[4] Finally, it is essential that all deposit copies submitted to the Copyright Office or Library of Congress under sections 407 or 408 of Title 17 retain full copyright protections and that government agencies do not use these deposits beyond what is presently allowed under the law.

---

[1] The Copyright Alliance supports comments submitted by its member organizations, including the International Association of Scientific, Technical & Medical Publishers (STM), but limits its comments to copyright issues that may be implicated by establishing a public-private consortium to curate the scientific data of DOE across the National Laboratory complex for use in AI models.

[2] United States, Executive Office of the President. America's AI Action Plan. The White House, July 2025, p. 8, www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf. (Explaining that it is "essential for the U.S. government to effectively address security risks to American AI companies, talent, intellectual property, and systems.)

[3] United States, Executive Office of the President. Launching the Genesis Mission. The White House. November 2025, Sec. 3, para. (v). https://www.whitehouse.gov/presidential-actions/2025/11/launching-the-genesis-mission/. (Explaining that accessing datasets must be "consistent with applicable law; applicable classification, privacy, and intellectual property protections.")

[4] See Appendix A.

Copyright Supports and Sustains the U.S. Scientific Innovation Ecosystem

Copyright law forms the backbone of the publishing ecosystem, ensuring the creation and dissemination of peer-reviewed journal articles, textbooks, databases, software, newspapers and magazines, and countless other literary works that are essential to the progress of science and the arts.[5] Copyright also enables sustainable business models for publishers and platforms through licensing markets, which have become even more pronounced in the AI era. Indeed, the proliferation of licensing agreements between AI developers and copyright owners in recent months underscores the immense value copyrighted works provide and the vibrant ecosystem they power.[6]

The U.S. continues to be the world leader in IP laws, which form the foundation of America's vast cultural influence, sustain America's leading role in scientific and medical advancements, and cement America's standing as the world's leading economy. Our IP industries support American dominance in scientific and technological innovation, and any partnerships, collaborations, reports, or policies that result from this RFI must take into account the effect any actions or policies may have on copyright industries' contributions to the economy and job creation. The Trump Administration has long been a proponent of strong intellectual property rights and an ally to the scientific publishing community, and any policies emanating from this RFI must be consistent with President Trump's vision for AI and IP.[7]

A Public-Private Consortium Must Recognize and Respect Established Copyright Laws

The best way to promote AI dominance and enable, not hinder, innovation is through the promotion of copyright law compliance that is based on the following principles.

---

[5] The publishing industry is responsible for over $30 billion of revenue annually in the U.S. alone. StatShot Annual Report, Association of American Publishers (Aug. 26, 2025); https://publishers.org/data-and-statistics/industry-statistics/.

[6] *Bringing Law and Order to the AI Wild West*, Keith Kupferschmid, COPYRIGHT ALLIANCE (Dec. 16, 2025), https://copyrightalliance.org/law-and-order-ai-wild-west/.

[7] See Appendix B for a summary of the numerous documents and orders issued by President Trump and his Administration that promote strong IP rights.

- Guidelines surrounding AI training methods using sensitive or proprietary data must include clear distinctions between purely government-generated scientific data and grantees' copyright protected materials.

- In most cases, licensing should be required for use of copyrighted works in AI training datasets.

- AI tools must not undermine markets for copyrighted works, including scientific publications and educational materials.

- Any regulatory carve-outs (including text and data mining exceptions) that would exempt AI companies from copyright compliance should be opposed.

- Data curation should be subject to best practices, including mandatory copyright clearance review before data inclusion.

An existing public-private consortium framework that the DOE should look to as it develops policy is the National Institute of Health's Generalist Repository Ecosystem Initiative (GREI), which (i) develops collaborative approaches for data management and sharing through the inclusion of generalist repositories in the NIH data ecosystem, and (ii) enables search and discovery of NIH-funded data in generalist repositories.[8] Through the implementation of best practices, including support for a wide range of copyright licenses and the promotion of transparency, the GREI model is one that respects existing copyright protections while ensuring authorized access and scientific advancement.

The Use of Fine-Tuning and Cloud Technologies Must be Accompanied by Copyright Protection Measures

To the extent fine-tuning of general-purpose AI models is necessary, guidelines must ensure that copyright protected materials provided by private partners are licensed or otherwise not used without authorization. Existing or already-planned Application Programming Interfaces (APIs) must be used in way that ensures the APIs don't retrieve or reproduce copyrighted works without

---

[8] National Institutes of Health, Office of Data Science Strategy. "Generalist Repository Ecosystem Initiative." Data Science at NIH. Last modified December 30, 2025. https://datascience.nih.gov/data-ecosystem/generalist-repository-ecosystem-initiative.

permission. Thus, when fine-tuning models, measures should be implemented that distinguish between purely government-generated data and licensed commercial materials. Additionally, while the use of cloud technologies may be preferable to ensure accelerated innovation in discovery science and engineering for new energy technologies, any models or datasets stored through cloud technologies must be accompanied by measures that safeguard against unauthorized access and support rights management via compliance with the Digital Millennium Copyright Act's (DMCA) anti-circumvention provisions.

Conclusion

Any policy reforms aimed at structuring partnerships for curating scientific data for use in AI models should explicitly preserve and affirm copyright protection for private sector works used for AI training as the baseline, and any new exemptions should be rejected. Indeed, maintaining strong copyright standards to protect U.S. publishers and platforms globally will help sustain U.S. leadership in scientific research and AI.

Respectfully Submitted,

Keith Kupferschmid
CEO
Copyright Alliance
1331 F Street, NW, Suite 950
Washington, D.C. 20004

January 14, 2026

APPENDIX A

Materials Created by Third Parties that Retain IP Rights Must be Distinguished from Purely Government-generated Scientific Data

The RFI asks questions about how to best structure partnerships for curating scientific data for use in AI models and for developing self-improving AI models for science and engineering using DOE's data. Ensuring that any AI policy developed pursuant to the RFI are compliant and consistent with U.S. IP laws requires that "data" be defined in a way that distinguishes between purely government-generated scientific data and any materials created by third parties that retain IP rights. To the extent the government uses copyright protected material for training that has been produced by grantees or other private sector entities, it should be made clear that such material retains copyright protection and cannot be used in a way that violates the copyright owner's rights.

The DOE definition of data is currently written broadly to mean "recorded information," and it explicitly includes (by reference) written materials and software.[9] It is essential that the definition of "data" must not be construed in a way that sweeps in copyright protected works generated by grantees or other private sector rightsholders. Licenses claimed by DOE for data for which a grantee retains copyright is broader than a federal purpose license, as it provides "a paid-up nonexclusive, irrevocable worldwide license" without any requirement that that license be applied for federal purposes. Such licenses must be subject to established copyright laws and not misconstrued to allow DOE to use protected works without securing appropriate permissions and licenses from private sector copyright owner grantees.

---

[9] 79 FR 76024, Dec. 19, 2014, as amended at 87 FR 15320, Mar. 18, 2022

APPENDIX B

- President Trump's Executive Order (EO) on AI makes clear that AI policies must "promote *human* flourishing, economic competitiveness, and national security."[10] (emphasis added).

- The Trump Administration's AI Action Plan explains that "[h]igh-quality data has become a national strategic asset."[11] There is no question that *human-created* copyrighted works produced by America's creative industries have become the backbone of the high-quality "data" AI companies want to use to train or augment AI models, and they are unequivocally a national strategic asset that should be protected with the same vehemence that patented AI chip technology and other forms of AI-related intellectual property are protected. In the Action plan the President also stresses how "essential [it is] for the U.S. government to effectively address security risks to American AI companies, talent, *intellectual property*, and systems."[12] (emphasis added).

- The Trump Administration's Genesis Mission EO states that the "Secretary [of Energy] shall establish and operate the American Science and Security Platform (Platform) to serve as the infrastructure for the Mission with the purpose of providing, in an integrated manner and to the maximum extent practicable and consistent with law" including by secure[ing] access to appropriate datasets, including proprietary, federally curated, and open scientific datasets, in addition to synthetic data generated through DOE computing resources, *consistent with applicable law; applicable classification, privacy, and intellectual property protections*; and Federal data-access and data-management standards.[13] (emphasis added)

---

[10] United States, Executive Office of the President. Removing Barriers to American Leadership in Artificial Intelligence. The White House, January 2025, Sec. 3, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/).

[11] *See* America's AI Action Plan, *supra* note 2.

[12] *Id*.

[13] *See* Launching the Genesis Mission, *supra* note 3.

- The Trump Administration's EO on Ensuring a National Policy Framework for Artificial Intelligence proposes any AI "framework should also ensure that children are protected, censorship is prevented, *copyrights are respected*…"[14] (emphasis added)

---

[14] United States, Executive Office of the President. Ensuring a National Policy Framework for Artificial Intelligence. The White House. December 2025, Sec. 1. https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/