

US Copyright Office Releases Third Report on AI Copyrightability, and House Committee Advances AI Legislation — AI: The Washington Report

May 16, 2025 | Article | By **Bruce Sokler**, **Alexander Hecht**, **Christian Tamotsu Fjeld**, Matthew Tikhonovsky

VIEWPOINT TOPICS

- Artificial Intelligence
- AI: The Washington Report

- On May 9, the United States Copyright Office (Office) issued a pre-publication version of its third Report on Copyright and Artificial Intelligence (Report), which focuses on the question of the use of copyrighted materials to train AI models under the fair use doctrine.
 - The 107-page Report does not draw broad conclusions, but it does suggest that the commercial use of copyrighted information to train AI models may not be protected under the fair use doctrine and “may infringe one or more rights.” According to the Report, however, the extent to which the use of copyrighted information is fair is highly context-specific and “depend[s] on what works were used, from what source, for what purpose, and with what controls on the outputs.”
 - The Report is nonbinding and comes as numerous courts are wrestling with the same question in lawsuits brought by the creative industry – cases that may bring additional clarity this year.
 - In Congress, on May 14, House Republicans advanced to the House floor a reconciliation package that includes AI measures. The proposed legislation would ban states from passing AI regulations for a 10-year period and appropriate \$500 million for modernizing federal IT systems with AI.
-

On May 9, the United States Copyright Office issued a pre-publication version of its third **Report** on Copyright and Artificial Intelligence (Report), which focuses on the question of the use of copyrighted materials to train AI models under the fair use doctrine. The 107-page Report does not draw broad conclusions, but it does suggest that the commercial use of copyrighted information to train AI models may not be protected under the fair use doctrine. According to the Report, however, the extent to which the use of copyrighted information is fair is highly context-specific and “depend[s] on what works were used, from what source, for what purpose, and with what controls on the outputs.” The Office spends much of the Report evaluating the factors that could weigh in favor of or against fair use for generative AI models and training.

In Congress, on May 14, House Republicans advanced to the House floor a reconciliation **package** that includes AI measures. The proposed legislation would ban states from passing AI regulations for a 10-year period and appropriate \$500 million for modernizing federal IT systems with AI.

Debate Around AI Models and Fair Use

The Copyright Office's Report acknowledges the “intense debate” around the question of the use of copyrighted materials to train AI models. “Dozens of lawsuits are pending in the United States, focusing on the application of copyright’s fair use doctrine,” while lawmakers around the world have considered legislation to “remove barriers or impose restrictions” on AI models. Just last week, the UK Parliament **rejected** a copyright measure for AI companies in its data bill.

The Report surveys the arguments for and against the use of copyrighted information to train AI models. On the one hand, the Office acknowledges that “some warn that requiring AI companies to license copyrighted works would throttle a transformative technology, because it is not practically possible to obtain licenses for the volume and diversity of content necessary to power cutting-edge systems.” On the other hand, “others fear that unlicensed training will corrode the creative ecosystem, with artists’ entire bodies of works used against their will to produce content that competes with them in the marketplace.”

Report’s Analysis of Gen AI

Early on, the Report draws a distinction between the use of copyrighted information by academics and nonprofits and the use of such information by commercial AI companies. “When a model is deployed for purposes such as analysis or research — the types of uses that are critical to international

competitiveness — the outputs are unlikely to substitute for expressive works used in training,” according to the Report. “But the making commercial use of vast troves of copyrighted works to produce expressive content that competes with them in existing markets, especially where this is accomplished through illegal access, goes beyond established fair use boundaries.”

For commercial use, the Office concludes that “creating and deploying a generative AI system using copyright-protected material involves multiple acts that, absent a license or other defense, may infringe one or more rights” and “clearly implicate the right of reproduction.”

Fair Use

After arguing that some commercial uses of copyrighted AI information “may constitute prima facie infringement,” the Office observes that the “primary defense available is fair use.”

The Report outlines four factors that should be considered in such a defense, which is highly case-specific:

1. The purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes
2. The nature of the copyrighted work
3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole
4. The effect of the use upon the potential market for or value of the copyrighted work

For the first factor – the purpose and character of the use – the Report notes that there is a spectrum of uses. On the one hand, when a model “generate[s] outputs that are substantially similar to copyrighted works in the dataset,” “it is hard to see the use as transformative.” But, on the other hand, “where a model is trained on specific types of works in order to produce content that shares the purpose of appealing to a particular audience,” that use could be “modestly transformative.”

“[W]hile it is important to identify the specific act of copying during development, compiling a dataset or training alone is rarely the ultimate purpose. Fair use must also be evaluated in the context of the overall use,” according to the Report.

For the second factor – the nature of the copyrighted work – the Office observes that “some works are closer to the core of intended copyright protection than others.” “The use of more creative or expressive works (such as novels, movies, art, or music) is less likely to be fair use than use of factual or functional works (such as computer code),” according to the Report. The Office recognizes that “generative AI models are regularly trained on a variety of works — both expressive and functional, published as well as unpublished,” which means the fair use determination will “vary depending on the model and works at issue.”

For the third factor – the amount and substantiality of the portion used in relation to the copyrighted work as a whole – the Office again observes that the analysis is fact dependent. “Copying even a small portion of a work may weigh against fair use where it is the ‘heart’ of the work,” according to the Report. But for Generative AI models, “downloading works, curating them into a training dataset, and training on that dataset generally involve using all or substantially all of those works. Such wholesale taking ordinarily weighs against fair use.” But other factors that could weigh in favor of or against fair use include not only how much of each work is used but also “the reasonableness of the amount in light of the purpose of the use and the amount made accessible to the public.”

And for the fourth factor – the effect of the use upon the potential market for or value of the copyrighted work – the Office observes that it is “undoubtedly the single most important element of fair use.” The Report focuses on three ways that generative AI can affect the market for copyrighted works, including “through lost sales, market dilution, and lost licensing opportunities,” which all “shift the fair use balance.”

- **Loss of Sales.** The Report identifies numerous instances “where the works in generative AI training can lead to a loss of sales.” It focuses on training for commercial generative AI models. “Where the content of [training] datasets is copyrightable, or the datasets themselves evince human selection and arrangement of data, and the datasets are primarily or solely targeted at AI training, widespread unlicensed use would likely cause market harm,” according to the Office.
- **Market Dilution.** Market dilution occurs when “a generative AI model’s outputs, even if not substantially similar to a specific copyrighted work, compete in the market for that type of work.” The Office acknowledges that the market dilution theory is “uncharted territory” and has not been tested in courts,” but the Office observes that “the speed and scale at which AI systems generate content pose a serious risk of diluting markets for works of the same kind as in their training data.”
- **Lost Licensing Opportunities.** “Lost revenue in actual or potential licensing markets can also be an element of market harm,” according to the Report.

The Report acknowledges that “it is for the courts to weigh the statutory factors together in light of the purposes of copyright, with no mechanical computation or easy formula.” The Office cautions that “as generative AI involves a spectrum of uses and impacts, it is not possible to prejudge litigation outcomes.”

The Office “expects that some uses of copyrighted works for generative AI will qualify as fair use, and some will not.”

Licensing

The Report spends its final section overviews licensing for AI training as a potential solution for the copyright issues that generative AI models may raise. The Office notes that “voluntary licensing may be workable, at least in certain contexts,” but it also has its challenges, such as scaling “for all AI training needs.” However, the Office contends that “a compulsory licensing regime for AI training would have significant disadvantages.” Premature adoption of a compulsory licensing regime “risks stifling the development of flexible and creative market-based solutions.”

The Report recognizes that “the current licensing market may be distorted by the unsettled legal questions about fair use” and different companies’ varying uses of AI and approaches to licensing. But “as courts begin to resolve pending cases, greater legal clarity may lead to greater collaboration on technical and market-based solutions” for licensing.

Controversy Around the Report

The Report itself addresses a controversial question, and its publication was surrounded by controversy. Over the weekend after the Report’s publication, President Trump, following his firing of the Librarian of Congress, dismissed the Register of Copyright; the last time a new President had done so was in 1861. It is unusual for the Copyright Office to release a “pre-publication version” of a report, and it is possible the Report may never be finalized.

The Office explained its decision to release a pre-publication version “in response to congressional inquiries and expressions of interest from stakeholders. A final version will be published in the near future, without any substantive changes expected in the analysis or conclusions.”

House Republicans Advance AI Legislation in Reconciliation Package

On May 12, House Republicans in the Energy and Commerce (E&C) Committee unveiled AI measures as part of the budget reconciliation package. The proposed measures would:

- Appropriate \$500 million to the Department of Commerce “for the purpose of modernizing and securing federal information technology systems through the deployment of commercial artificial intelligence.”
- Prohibit any state or political subdivision from “[enforcing] any law or regulation regulating artificial intelligence models, artificial intelligence systems, or automated decision systems during the 10-year period beginning on the date of the enactment of this Act.”

On May 14, lawmakers on the E&C Committee voted 29-24 to advance the AI measures. During the mark-up session, House E&C Chair Rep. Brett Guthrie (R-KY) remarked, “Through investments to modernize the Department of Commerce, we can integrate AI systems to make the Department more secure and effective. And we’re implementing guardrails that protect against state level AI laws that could jeopardize technological leadership.”

During the mark-up session, House Democrats tried but failed to remove the 10-year moratorium on states enforcing AI laws. Rep. Frank Pallone (D-NJ) called the measure “an unprecedented giveaway to Big Tech.” Responding to Democrats’ concerns about inaction on federal AI legislation, Chair Guthrie remarked, “We know we need to have a national standard.”

The reconciliation package is headed to the House floor, where there could be further amendments, and then to the Senate and final conference. On May 15, Senator Ted Cruz (R-TX) announced his plans to introduce an AI bill in the Senate with a 10-year moratorium on state AI laws. He also said he will “soon release a new bill that creates a regulatory sandbox for AI ... that will remove barriers to AI adoption, prevent needless state over-regulation, and allow the AI supply chain to rapidly grow here in the US” The AI measures’ Republican support in the Republican-controlled Congress means that their path to passage is becoming increasingly clear.

We will continue to monitor, analyze, and issue reports on these developments. Please feel free to contact us if you have questions as to current practices or how to proceed.

Authors

Bruce Sokler

Bruce D. Sokler is a Mintz antitrust attorney. His antitrust experience includes litigation, class actions, government merger reviews and investigations, and cartel-related issues. Bruce focuses on the health care, communications, and retail industries, from start-ups to Fortune 100 companies.

Alexander Hecht, Executive Vice President & Director of Operations



Alex Hecht is a trusted attorney and policy strategist with over 20 years of experience advising clients across a broad range of industries on how to navigate complex policy environments. His strategic insight and hands-on experience in both legislative and regulatory arenas empower clients to advance their priorities with clarity and confidence in an evolving policy landscape.

Christian Tamotsu Fjeld, Senior Vice President



Christian Tamotsu Fjeld draws on two decades of Capitol Hill experience to support clients in building relationships, shaping policy, and engaging effectively with the federal government. His experience working with Congress and his insights help clients anticipate federal developments and advance their priorities with clarity and confidence.

Matthew Tikhonovsky

Matthew is a Mintz Senior Project Analyst based in Washington, DC.