The Computer & Internet Lawyer

Volume 43 ▲ Number 4 ▲ April 2026

Ronald L. Johnston, Arnold & Porter, Editor-in-Chief

The Open Questions in U.S. Generative AI Copyright Litigation

By Angela L. Dunning and Arminda B. Bepko

Important developments in the legal landscape of generative AI and copyright are coming. Dozens of copyright infringement lawsuits targeting the training and development of AI models—capable of generating text, images, video, music and more—are advancing toward dispositive rulings. The central issue remains whether training AI models using unlicensed copyrighted works is infringing or instead constitutes fair use under Section 107 of the U.S. Copyright Act.

Courts consider four factors in determining whether a particular use is fair:

(1) the purpose and character of the use;
(2) the nature of the copyrighted work;
(3) the amount and substantiality of the portion used; and
(4) the effect of the use upon the potential market for or value of the copyrighted work.

The authors are attorneys at Cleary Gottlieb Steen & Hamilton LLP. They may be contacted at adunning@cgsh.com and abepko@cgsh.com, respectively.

The thrust of this inquiry is whether the use is transformative—serving a different purpose or function from the original work—or merely usurps the market for the original by reproducing its protected expression. As courts establish legal frameworks for AI training and protection of AI-generated outputs, companies and boards should closely monitor developments to fully understand the risks and opportunities of AI implementation.

OVERVIEW OF AI COPYRIGHT LITIGATION

The first court to reach a substantive decision this year on fair use in the context of an AI-augmented platform was *Thomson Reuters Enterprise Center GmbH v. ROSS Intelligence Inc.*[1] Thomson Reuters sued ROSS Intelligence for allegedly copying headnotes from Westlaw, Thomson Reuters' legal research platform, to train ROSS's AI-based legal research platform. In February 2025, the district court sided with Thomson Reuters and rejected ROSS's fair use defense. Taking pains to note that ROSS's platform did not involve generative AI, the court found that factor four strongly favored Thomson Reuters because ROSS sought to

create a platform that directly competed with Westlaw and thereby harmed both the market for its legal-research services and a potential derivative market for data to train legal AI systems. However, on April 4, 2025, the district court certified two questions for interlocutory appeal to the U.S. Court of Appeals for the Third Circuit: (1) whether Westlaw headnotes constitute original expression protectable under the Copyright Act, and (2) whether ROSS's use was fair. In a subsequent memorandum opinion explaining its reasoning, the district court described these as "hard" issues presenting substantial grounds for disagreement.[2]

In June 2025, two judges in the U.S. District Court for the Northern District of California reached a different conclusion on fair use at summary judgment, finding that the use of copyrighted works to train generative AI models that do not substantially reproduce the content on which they were trained was a transformative fair use as a matter of law. In *Bartz v. Anthropic PBC*, Judge William Alsup held that Anthropic's use of copyrighted books to train Claude, Anthropic's large language model, was fair, calling the technology "spectacularly" transformative.[3] The court also rejected plaintiffs' contention that training LLMs on their books "will result in an explosion of works" that compete with theirs, finding that any competition arising from non-infringing outputs would not constitute cognizable market harm under the Copyright Act."[4]

However, Judge Alsup refused to grant summary judgment for Anthropic as to copies of books obtained from "pirate" websites and allegedly maintained in "a permanent, general-purpose resource even after deciding it would not use certain copies to train LLMs or would never use them again to do so."[5] The court expressed significant skepticism that such a use could be justified as transformative and fair and ordered that claim to trial. Shortly thereafter, the court certified a class consisting of book authors holding valid copyrights to works downloaded from certain specific "pirate" datasets.[6] Anthropic subsequently entered into a class wide settlement pursuant to which it agreed to pay class plaintiffs $1.5 billion, or an estimated $3,000 per book. The court granted preliminary approval for the settlement on October 17, 2025, and final approval remains pending.

Days after the summary judgment decision issued in Anthropic, Judge Vincent Chhabria granted summary judgment for Meta in *Kadrey v. Meta Platforms, Inc.* The court held that Meta's use of copyrighted books to train Llama was "highly transformative."[7] On factor four, Judge Chhabria rejected plaintiffs' harm theories as "clear losers,"[8] concluding that plaintiffs had failed to adduce any evidence that Llama's outputs reproduced plaintiffs' protected expression or had harmed sales of plaintiffs' books.[9]

Similarly, Judge Chhabria rejected arguments that Meta's training without permission diminished plaintiffs' ability to license their works for training, noting that such a market is "theoretical" and that the market for a transformative use is not one plaintiffs have the right to monopolize in any event.[10]

The court also rejected plaintiffs' argument that Meta's use of their books to train Llama could not be fair because it had acquired copies of those books from "pirate" websites in bad faith, finding that this factor did "not move the needle" given the transformativeness of the use and the lack of market harm.[11]

Finally, the court identified a possible "market dilution" theory of harm—where AI outputs flood markets and thereby harm sales of the human-authored works on which they were trained—but found plaintiffs had proffered no evidentiary support for such a theory.[12]

While no Circuit Court has yet decided AI fair use or addressed the reasoning applied by the Judges in the Anthropic and Meta cases, the U.S. Court of Appeals for the Second Circuit's 2025 opinion in *Romanova v. Amilus, Inc.*, may provide guidance. Judge Pierre N. Leval—an authority on the modern fair use doctrine and author of several seminal decisions on transformative fair use—authored the opinion, which includes a comprehensive discussion of fair use principles.[13] In particular, the discussion of factor four (market harm) emphasizes the classic market substitution injury—whether the new use is substitutive because it makes "the protected expression of the original" available to the public for a non-transformative purpose.[14] This decision, along with Judge Leval's earlier decision in *Authors Guild v. Google*,[15] should provide substantial guidance to the U.S. District Court for the Southern District of New York, which is scheduled to reach summary judgment on issues of infringement and fair use in *In re OpenAI, Inc., Copyright Infringement Litigation*[16] in August. Dispositive motions are also expected this year in *In re Google Generative AI Copyright Litigation*,[17] *Andersen v. Stability AI*,[18] and *Nazemian v. NVIDIA Corporation*.[19]

The critical question in all of these cases will be whether generative AI models promote the creation of new, original expression and serve a distinct purpose from the works on which they were trained or instead substitute for those works by making accessible their protected expression. For practitioners on both sides of these disputes, 2026 will be a defining year as courts chart the boundaries of fair use in the AI era.

REGULATORY GUIDANCE ON COPYRIGHT & AI

The U.S. Copyright Office has issued guidance on several AI-related intellectual property issues. The first installment of this guidance—on the use of digital technology to replicate an individual's voice or appearance—was released in July 2024.[20] In January 2025, it addressed copyrightability of AI-produced works, in particular how much human authorship is required for protection.[21] While applications to register AI generated outputs had previously been rejected on grounds that they lacked the necessary human authorship, the Office issued its first such registration in January 2025, for "A Single Piece of American Cheese."

While holding fast to its position that human authorship is required, the Copyright Office determined in this instance that the process of selecting, arranging and coordinating different iterations of AI-generated material and piecing them together to make a composite work involved the requisite level of human control necessary to satisfy that requirement.

How much human authorship is required for registration, and what that authorship can properly consist of, remain open issues for the courts to decide. The first such case is set to be decided this year in *Allen v. U.S. Copyright Office*, pending in the U.S. District Court for the District of Colorado.[22] It involves an author whose award-winning visual artwork, "Théâtre D'Opéra Spatial," was repeatedly rejected for registration by the Copyright Office for lack of human authorship.

The artist sued to overturn the rejection, arguing that his creative decisions in crafting more than 600 iterative prompts constitute sufficient originality and human authorship to warrant copyright protection, analogizing his process to the creative choices photographers make in composing images. Summary judgment will be fully briefed in early 2026, and the outcome of this decision could have significant import for any company or individual author seeking to protect works created with the benefit of generative AI tools.

In May 2025, the Copyright Office issued "prepublication" guidance addressing fair use in AI training,[23] noting that fair use outcomes will be highly fact-specific across different factors. In particular, training AI models can often be transformative (factor one)—particularly for research or non-substitutive uses—while uses aimed at generating substantially similar expressive outputs or occupying the same markets may be less so in its view.[24] Factor two (nature of the work) will weigh against fair use where "works involved are more expressive, or previously unpublished,"[25] and factor three (amount used) may weigh against fair use given wholesale copying, though it may favor fair use where such copying serves a transformative purpose and little protectable material is made accessible through outputs with effective guardrails.[26] While noting that this is "uncharted territory," the guidance also endorses a potential "market dilution" theory for factor four (market harm) similar to *Kadrey,* alongside traditional harms of lost sales and licensing opportunities for the works at issue.[27] The guidance is not controlling, however, as the question of when the training of generative AI models constitutes fair use is a highly case-specific one that will ultimately be determined by federal courts in the context of adjudicating actual disputes or by Congress, should it choose to step in.

Notes

1. Thomson Reuters Enterprise Center GmbH v. ROSS Intelligence Inc., 765 F. Supp. 3d 382 (D. Del. 2025), motion to certify appeal granted, No. 1:20-CV-613-SB, 2025 WL 1488015 (D. Del. May 23, 2025).
2. Thomson Reuters Enterprise Centre GmbH et al v. ROSS Intelligence Inc., No. 1:20-cv-00613-SB, Memorandum Opinion, ECF No. 804 at 1 (D. Del. May 23, 2025).
3. 787 F. Supp. 3d 1007, 1019, 1021 (N.D. Cal. 2025).
4. Id. at 1031-32.
5. Id. at 1014.
6. Bartz v. Anthropic PBC, 791 F. Supp. 3d 1038 (N.D. Cal. 2025).
7. Kadrey v. Meta Platforms, Inc., 788 F. Supp. 3d 1026, 1044 (N.D. Cal. 2025).
8. Id. at 1036.
9. Id. at 1051.
10. Id. at 1052.
11. Id. at 1058.
12. Id. at 1036 (N.D. Cal. June 25, 2025).
13. Romanova v. Amilus, Inc., 138 F.4th 104 (2d Cir. May 23, 2025).
14. Id. at n.9.
15. Authors Guild v. Google, 804 F.3d 202 (2d Cir. 2015).
16. In re OpenAI, Inc., Copyright Infringement Litigation, No. 25-md-3143-SHS (S.D. NY).
17. In re Google Generative AI Copyright Litigation, No. 23-cv-00201-WHO (N.D. Cal.).
18. Andersen v. Stability AI, No. 23-cv-03440-EKL (N.D. Cal.).
19. Nazemian v. NVIDIA Corporation, No. 23-cv-00201-WHO (N.D. Cal.).
20. See U.S. Copyright Office, "Copyright and Artificial Intelligence Part 1: Digital Replicas" (July 2024), available at https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf.
21. See U.S. Copyright Office, "Copyright and Artificial Intelligence Part 2: Copyrightability" (January 2025), available

at https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf.

22. Allen v. U.S. Copyright Office, No. 24-cv-02665-WJM (D. Colo.).
23. See U.S. Copyright Office, "Copyright and Artificial Intelligence Part 3: Generative AI Training (Pre-Publication Version)" (May 2025), available at https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf.
24. Id. at 45-46.
25. Id. at 54.
26. Id. at 59-60.
27. Id. at 65.