

UPDATE ON GENERATIVE AI LITIGATION

A FUTURE HANGING IN THE BALANCE

BY ELLIE K. VILENDRER



Illustration by Cemile Bingol via Getty Images

The age of generative artificial intelligence (GenAI) litigation began roughly five years ago, when media company Thomson Reuters sued AI research firm Ross Intelligence asserting that Ross had trained its AI tools using Westlaw’s legal summaries without first obtaining a license to do so.¹ Like a roller coaster, the wheels of that early lawsuit inexorably pulled the GenAI litigation train up the track, with more cars being added as it made its ascent. The downhill ride has been both fast and furious, with dozens of new cars being added to the ride and litigants jumping aboard from all sides.

These cases have encompassed virtually every aspect of GenAI, from the information-gathering process to the distribution and dissemination of information. They have asserted copyright, trademark, and privacy claims. They have been filed by authors, musicians, artists, news organizations, and others who believe they alone should have the right to exploit and monetize their creations.

Appropriation of prior works may implicate the methods used to obtain data, from scraping to wholesale copying of those works. It may raise both privacy and copyright issues. At the other end of the process, plaintiffs are alleging that materials generated by AI infringe their intellectual property rights and deprive them of revenue to which they are entitled.

At the heart of many of the debates about AI’s impact on creative fields are questions of fair use, namely, whether AI models trained on copyrighted works are covered, at least in the United States, by that doctrine. As a practical matter, most GenAI models have been trained by scraping large amounts of text from various sites across the web.

AI companies have argued that creating AI tools is a legitimate reason to use copyrighted materials without getting consent or paying compensation to rights holders. The fair use doctrine invokes four measures when evaluating whether a work is “transformative” or simply a copy: the purpose and character of the work, the nature of the work, the amount taken from the original

work, and the effect of the new work on a potential market. Most AI companies have asserted that their use of existing works is sufficiently transformative to constitute fair use.

In 2023, the U.S. Copyright Office undertook an initiative to examine “the copyright law and policy issues raised by artificial intelligence (AI).”² Part 1 of the three-part report, on digital replicas, was published on July 31, 2024. Part 2, published on January 29, 2025, addressed the copyrightability of outputs created using generative AI. Part 3 of the report will address the legal implications of training AI models on copyrighted works, licensing considerations, and the allocation of potential liability.

At the start of 2025, the GenAI lawsuit ride is operating at full speed with new cars being added to an ever-growing track. Earlier cases have been consolidated and amended; new cases have been filed. Both plaintiffs and defendants have refined and strengthened their arguments, learning from prior cases how to better present their positions.

The next year should bring closure to some important issues, potentially ending the free ride enjoyed by AI companies for certain activities. Here we review the spiraling path these cases have taken and explore their implications for future GenAI litigation.

TRAINING GENAI

At their heart, most training GenAI cases deal with the use of existing works—images, music, and writings—to train AI tools.³ In its precedential lawsuit, Thompson Reuters asserted that Ross misappropriated Westlaw headnotes for purposes of training a competitive legal AI tool without having to spend the resources, creative energy, and time to create it itself. Nearly five years later, this case has reached the trial stage. It is the first case to test the theory that training an AI’s learning model with copyrighted material should be considered a fair use.

What may have been the first GenAI class-action lawsuit, *Doe v. Github*, was filed in 2022 by Github contributors.⁴

Their lawsuit alleged that Github and Open AI used the plaintiffs’ works to train the Copilot and Codex AI tools in violation of their open-source license agreements and that the defendants removed copyright management information (CMI) from their code in violation of the Digital Millennium Copyright Act (DMCA).⁵

Section 1202(b) of the DMCA provides that one cannot, without authority, (1) “intentionally remove or alter any” CMI, (2) “distribute . . . [CMI] knowing that the [CMI] has been removed or altered,” or (3) “distribute . . . copies of works . . . knowing that [CMI] has been removed or altered” while “knowing, or . . . having reasonable grounds to know, that it will induce, enable, facilitate, or conceal” infringement.⁶

The DMCA is a statute with sharp teeth. Section 1203 grants the court broad powers to issue civil remedies for Section 1202 violations, including ordering injunctions, product impounding and destruction, and awarding actual or statutory damages, costs, and attorney fees.

In *Doe v. Github*, the plaintiffs claimed that, when prompted to generate software code, Copilot includes unique aspects of the plaintiffs’ code in its outputs, and in removing CMI, the defendants failed to prevent users of products from making noninfringing use of the product. Moreover, plaintiffs alleged that defendants’ false description of the source of Copilot’s output facilitated or concealed infringement by defendants and Copilot users.

A 2023 lawsuit by a group of eight music publishers against Anthropic PBC, an “AI safety and research company,” claimed that the AI company used the plaintiffs’ copyrighted lyrics (along with “vast amounts of text copied from the internet, totaling billions or trillions of words”) to train its signature product—a series of generative AI conversational interface models referred to as “Claude.”⁷

Two cases against Stability AI—*Andersen v. Stability AI Ltd.*⁸ and *Getty Images v. Stability AI Ltd.*⁹—were filed at the beginning of 2023. Both actions

involve copyright infringement claims based on the defendant's appropriation of protected images to train AI image generators. The latter case, which alleges Stability AI has copied more than 12 million photographs from Getty Images' collection, also includes trademark infringement claims arising from the output generated by the accused technology, which is at times "lower quality and ranges from the bizarre to the grotesque."¹⁰ Both cases are still in progress.

Dozens of well-known authors filed lawsuits against AI companies in 2023, alleging that those companies used copyrighted works, without permission, to train their AI language models. *Kadrey v. Meta*, filed in July 2023,¹¹ pitted a list of writers, including comedian Sarah Silverman and journalist Ta-Nehesi Coates, against the social media giant. After other authors withdrew from the case, Silverman's action was consolidated with a pending case brought by visual artists against Google.

In *Authors Guild v. OpenAI*, filed in September 2023,¹² the Authors Guild, John Grisham, and 16 other authors filed a class action suit on behalf of a class of fiction writers whose works have been used to train GPT. The suit alleged that OpenAI had infringed on the authors' copyrights by using their written works to train its models to output human-seeming text responses to users' prompts and queries. The action, originally filed as three separate cases, was consolidated into a single case and remains open.

The New York Times (NYT) sued Microsoft and OpenAI the last week of 2023, claiming that the companies had copied and used millions of the publisher's works, including copyrighted news articles, in-depth investigations, opinion pieces, reviews, and how-to guides, without a license, for purposes of training ChatGPT.¹³ The defendants, according to the complaint, engaged in direct, vicarious, and contributory copyright infringement and removed CMI in violation of the DMCA. The lawsuit followed purported attempts to reach a settlement, which would have avoided costly litigation.

According to the complaint, "The Times has attempted to reach a negotiated agreement with Defendants," with the goal of ensuring "it received fair value for the use of its content, facilitate the continuation of a healthy news ecosystem, and help develop GenAI technology in a responsible way that benefits society and supports a well-informed public."¹⁴ Noting defendants' reliance on "fair use," the complaint argues that "there is nothing 'transformative' about using The Times's content without payment to create products that substitute for The Times and steal audiences away from it."¹⁵ The action seeks "billions of dollars in statutory and

**DIRECT COPYING
OF COPYRIGHTED
WORKS FOR
TRAINING
PURPOSES MAY BE
GROUND ZERO IN
GENAI LITIGATION
THIS NEXT YEAR.
AI COMPANIES
DON'T ENGAGE IN
DIRECT COPYING.**

actual damages that they owe for the unlawful copying and use of The Times's uniquely valuable works."¹⁶

The NYT case was later consolidated with lawsuits filed against the defendants by two other news organizations: *Daily News v. Microsoft Corp.* and *The Center for Investigative Reporting v. OpenAI*. In November 2024, the court denied a motion by OpenAI to compel the NYT to produce evidence on its employees' use of generative AI, ruling that such evidence had no bearing on OpenAI's fair use defense.

On January 14, 2025, the news organizations were in federal court arguing against OpenAI's motion to dismiss

their case. The defense argued that its use of their articles was fair use and transformative; the plaintiffs argued that ChatGPT's taking of copyrighted works on a massive scale without license or payment was in essence a free ride, setting it up to directly compete with the news organizations as a substitute for the publishers' original work.

None of these cases has yet been resolved, and at least a dozen more GenAI lawsuits await resolution. Training is clearly an essential and ubiquitous element of GenAI tools, so the stakes are high for both sides. Direct copying of copyrighted works for training purposes may be Ground Zero in GenAI litigation this next year. AI companies contend they don't actually engage in direct copying; their computers "read" text, evaluate it, and then discard it. Any residual content is sufficiently remote from the original material to qualify the resulting work as "transformative." If their activity—scraping data and reading text—is ultimately found to be illegal, it could completely upend the industry.

Highlighting a breadth of possible claims in GenAI litigation, Match Group LLC, owner of Match.com, Tinder, and dozens of other dating sites, filed suit in Dallas County, Texas, against a startup alleging that it used AI romantic prospects to stand in for actual humans. The plaintiffs claimed that CupidBot.ai used AI to automate the process of browsing user profiles, matching clients with other users—all of whom believe they are communicating with actual human beings—"and deceptively striking up conversations and making dates." Their lawsuit asserts causes of action for trademark infringement, tortious interference with contract, violation of the Computer Fraud and Abuse Act, harmful access by computer, and breach of contract.¹⁷

Lawsuits also have been filed against AI companies for defamation.¹⁸ Public figures such as law professor Jonathan Turley and Australian mayor Brian Hood have raised the flag about the harm to reputation caused by false statements made against them by GenAI. False outputs, known as

“hallucinations”—which have taken the form of fabricated legal citations, nonexistent news articles, imaginary scientific data and statistics, inaccurate biographical information, and false historical or current events—pose a reliability risk that could lead to legal claims in the future, not just against the creators of AI, but also against downstream users.

2024 IN REVIEW

Last year saw significant movement on cases that had been filed in prior years, as well as the commencement of several new matters.

Two online news organizations filed suit in early 2024 against OpenAI. In *Raw Story Media v. OpenAI* and *The Intercept Media, Inc. v. OpenAI*, the plaintiffs claimed violation of Section 1202(b) of the DMCA based on the defendant’s alleged removal of CMI from works that were used to train ChatGPT.¹⁹ The court granted defense motions to dismiss the cases, finding that the plaintiffs lacked standing to pursue injunctive relief or damages. Citing *Spokeo, Inc. v. Robins*,²⁰ the court wrote that “Article III standing requires a concrete injury even in the context of a statutory violation.”²¹ In this case, the plaintiffs “have not alleged any *actual* adverse effects stemming from this alleged DMCA violation.”²² The plaintiffs then sought leave to amend their complaints.

In December, the US District Court for the Southern District of New York ruled partially in favor of The Intercept. The court dismissed, with prejudice, The Intercept’s claims under 17 U.S.C. § 1202(b)(3)—barring distribution, import for distribution, or public performance of works, copies of works, or phonorecords with knowledge that CMI has been removed or altered without authority of the copyright owner or the law—but it allowed the plaintiff’s claim under 17 U.S.C. § 1202(b)(1)—for removal of CMI from articles used to train OpenAI’s large language models (LLMs)—to proceed past the motion-to-dismiss stage.

The decision could serve as a precedent for other plaintiffs challenging the unauthorized use of their works by

AI developers. OpenAI told the court that it would try to consolidate the eight copyright infringement and DMCA lawsuits now pending against it into a single multidistrict litigation in the Northern District of California.

Also in the Southern District of New York, in October 2024 two different news publishers, Dow Jones & Company, Inc., and NYP Holdings, Inc., filed an action against Perplexity AI, Inc., based on use of its copyrighted works for a “retrieval-augmented generation,” or RAG, database. In *Dow Jones v. Perplexity*,²³ the plaintiffs took issue with the appropriation of its copyrighted works for use as part of a search engine that incorporates vast numbers of webpages and provides that information to previously developed and trained LLMs. The defendant, according to the complaint, used these models to repackage copyrighted works into verbatim or near-verbatim responses to user prompts.

Lawsuits brought in 2024 against AI music generators asserted that the defendants had used copyrighted recordings to train their models. Two lawsuits filed in 2024 pitted record labels against AI music generators. *UMG Recordings, Inc. v. Suno, Inc.* and *UMG Recordings, Inc. v. Udio* both allege direct copyright infringement based on defendants’ use of music recordings to train their AI models.²⁴ The lawsuits contend that the end product mimics unique features of the plaintiffs’ works. In response, the defendants argue that copying the works constitutes fair use, as an “intermediate” step.

Another case brought under the DMCA, *Vacker v. ElevenLabs*, asserts claims on behalf of a group of voice actors who allege that removal or alteration of CMI from their copyrighted works violated Section 1202.²⁵ Like the charges asserted against OpenAI by actress Scarlett Johansen,²⁶ the voice-actor plaintiffs in *Vacker* also allege that the defendant misappropriated their voices and likenesses by creating voice clones that it used to attract millions of users, generating significant revenue. According to the complaint, these voice clones mimicked their distinctive vocal

timbres, accents, intonation, pacing, vocal mannerisms, and speaking styles, delivering a synthetic professional narration that friends and family would recognize as their voices. With so much riding on the DMCA claim regarding removal or alteration of CMI, both sides in this case—as well as many other litigants—will be watching developments in *Doe v. Github*.

2025: WHAT TO EXPECT

Despite the whirlwind associated with boarding the AI roller coaster, GenAI cases have actually been moving at a snail’s pace. The tension between creators’ rights and the advent of world-changing technology appears to be forcing courts to tread extremely cautiously. Those delays could ultimately end up harming plaintiffs, as AI rapidly becomes a *de facto* part of our everyday life. It may soon be impossible to unravel the intricate web spun around us by the AI industry.

Another area ripe for litigation or regulation in the next year is developer liability for misuse by third parties. Earlier this year, the Federal Trade Commission (FTC) provided a preview of its approach to AI under the new administration. On January 9 during a panel at the 2025 CES event in Las Vegas, two sitting FTC commissioners discussed the FTC’s approach to AI. Both commissioners generally agreed that a developer who develops general-purpose AI tools that are misused by third parties, without additional involvement or knowledge by the developer, should not be subject to liability under the FTC Act. They disagreed, however, on where the line should be drawn.²⁷

On December 18, 2024, the FTC announced settlement of *In re Rytr LLC*.²⁸ Rytr is an internet service that uses GenAI to produce written content for subscribers for “Use Cases,” one of which is for testimonials and reviews. In the settlement, the FTC majority argued that the developer of a GenAI tool that could be used to generate large numbers of deceptive reviews could be held liable, based on the facts of that case. The dissent argued that “[t]reating as categorically illegal a generative

AI tool merely because of the possibility that someone might use it for fraud . . . threatens to turn honest innovators into lawbreakers and risks strangling a potentially revolutionary technology in its cradle.”²⁹

Policy under the new administration is tending to favor big business, with President Donald Trump—who revoked a 2023 executive order signed by former President Joe Biden that sought to reduce the potential risks AI poses to consumers, workers, and national security—saying he would “support AI development rooted in free speech and human flourishing.”³⁰

The American Bar Association is also contributing to the development of GenAI policy. On February 3 of this year, the ABA adopted Resolution 501—urging enactment of federal legislation protecting an individual’s right to authorize or prevent any use of their voice, visual likeness, and/or image in a realistic computer-generated electronic representation. The resolution, sponsored by the Section of Intellectual Property Law, stated that such legislation should (a) include strong safeguards to ensure the legislation’s compatibility with the First Amendment and (b) address the right of publicity and the right of privacy under state, territorial, or tribal law; technological innovation and creations; and potential third-party liability.³¹

We can expect to see significant judicial action involving GenAI over the next few months. *Thompson Reuters*, the original AI case, is expected to finally go to trial this year. Assuming the parties do not settle on the proverbial courthouse steps, that decision could establish precedent for all current and future GenAI cases. If the court rules in favor of the plaintiff, AI companies may be forced to decide whether to negotiate separate licenses with individual copyright owners or face a much stronger wave of lawsuits. We may see the creation of global licensing agencies—along the lines of ASCAP and BMI for music—to manage the use of copyrighted content.

An interlocutory appeal in the case of *Doe v. Github* also will be heard soon. That case, which focuses on the DMCA’s prohibition against altering

or removing CMI from copyrighted works, could determine whether Section 1202(b) claims will be allowed to move forward or must be excised from AI complaints. All eyes are sure to be trained on these and other cases as the GenAI roller coaster begins its next full year of operation.

The last year was a pivotal one for GenAI litigation. Plaintiffs’ claims continued to focus on infringement and DMCA violations; defenses generally relied on fair use and standing issues. Cases were consolidated, moved to other venues, and amended. Courts took seriously the economic and societal implications of GenAI, requiring parties on both sides to build strong supports for the carnival ride upon which their cars would be rolling at lightning speed.

This year we should expect to see some important decisions. Are data scraping and other content appropriation transformative? How and when is CMI alteration harmful to copyright owners? At what point is a new work separate and discrete from a prior work? Can GenAI be trained to do the work for which it was created without trespassing on sacred ground?

Ellie K. Vilendrer, CIPP/US, CIPT, is a mediator and arbitrator with Signature Resolution. She is the budget officer of the ABA Section of Dispute Resolution and chair of the Orange County Bar Association Legislative Resolutions Committee. She has extensive experience serving technology companies ranging from startup to publicly traded, in roles including privacy counsel, intellectual property counsel, and general counsel. In 2013, she co-founded a tech company that pioneered an artificial intelligence-powered platform that provided predictive insights for app features and marketing campaigns.

ENDNOTES

1. Complaint, *Thompson Reuters Enter. Ctr. GMBH v. Ross Intel. Inc.*, Case 1:20-cv-00613-SB (D. Del. May 6, 2020); Mem. Op. (D. Del. Sep. 25, 2023, 694 F.Supp.3d 467); Mem. Op. (D. Del. Apr. 26, 2022); Mem. Op. Feb. 11, 2025; Order granting motion

for interlocutory appeal and stay pending appeal (D. Del. Apr. 4, 2025).

2. *Copyright and Artificial Intelligence*, US COPYRIGHT OFF., <https://www.copyright.gov/ai/>.

3. Some AI models are being trained, at least in part, on AI-generated data. This can lead to a phenomenon known as “model collapse”—where AI models degrade in quality over generations due to errors and biases compounding in the dataset.

4. *Doe v. Github, Inc.*, Case No. 4:22-cv-06823-JST, 2024 U.S. Dist. LEXIS 175951 (N.D. Cal. Sept. 27, 2024).

5. Digital Millennium Copyright Act (DMCA), Pub. L. No. 105-304, 112 Stat. 2860 (1998).

6. *Id.* § 1202(b), 112 Stat. at 2872.

7. Complaint, *Concord Music Grp. v. Anthropic PBC*, Case No. 3:23-cv-01092 (M.D. Tenn.), *transferred to* Case No. 5:24-cv-03811-EKL (N.D. Cal. June 24, 2024).

8. 700 F. Supp. 3d 853 (N.D. Cal. 2023); Order granting in part and denying in part motions to dismiss first amended complaint, 23-cv-00201-WHO (N.D. Cal. Aug. 12, 2024).

9. Complaint, *Getty Images (US), Inc. v. Stability AI, Inc.*, Case No. 1:23-cv-00135-GBW (D. Del. Feb. 3, 2023), <https://docs.justia.com/cases/federal/district-courts/delaware/dedce/1:2023cv00135/81407/1>; Request for Oral Argument (D. Del. Aug. 26, 2024).

10. *Id.* at 4.

11. Complaint, *Kadrey v. Meta Platforms, Inc.*, Case No. 23-cv-03417-VC (N.D. Cal. July 7, 2023).

12. Case 1:23-cv-08292-SHS (S.D.N.Y. Dec. 4, 2023).

13. Complaint, *N.Y. Times Co. v. Microsoft Corp.*, Case No. 1:23-cv-11195 (S.D.N.Y. Dec. 27, 2023).

14. *Id.* at 3.

15. *Id.* at 4.

16. *Id.*

17. Complaint, *Match Grp. LLC v. CupidBot AI, Inc.* (D.C. Tex. Mar. 12, 2024).

18. *See, e.g.*, Complaint, *Jeffery Battle v. Microsoft Corp.*, Case No. 1:23-cv-01822 (D. Md. July 7, 2023); Complaint, *Walters v. OpenAI, L.L.C.*, Case No. 1:23-cv-03122 (N.D. Ga. July 14, 2023).

19. Complaint, *Raw Story Media v. OpenAI*, Case No. 24-cv-01514 (S.D.N.Y. Apr.

10, 2024); Complaint, *The Intercept Media, Inc. v. OpenAI*, Case No. 24-CV-1515 (S.D.N.Y. Feb. 28, 2024).

20. 578 U.S. 330, 341 (2016).

21. *Raw Story Media v. OpenAI, LLC*, Case No. 1:24-CV-01514-CM, 2024 U.S. Dist. LEXIS 204101, at *6 (S.D.N.Y. Nov. 7, 2024).

22. *Id.* at *10.

23. Complaint, *Dow Jones & Co., Inc. v. Perplexity AI, Inc.*, No. 24-cv-7984-KPF (S.D.N.Y. Dec. 11, 2024).

24. Complaint, *UMG Recordings, Inc. v. Suno, Inc.*, Case No. 1:2024cv11611 (D. Mass. June 24, 2024); Complaint, *UMG Recordings, Inc. v. Uncharted Labs, Inc. (Udio)*, Case No. 1:24-cv-04777-AKH (S.D.N.Y. June 25, 2024).

25. Complaint, *Vacker v. ElevenLabs, Inc.*, Case No. 1:2024cv00987 (D. Del. Aug. 29, 2024).

26. Bobby Allyn, *Scarlett Johansson Says She Is “Shocked, Angered” over New ChatGPT Voice*, NPR (May 20, 2024), <https://www.npr.org/2024/05/20/1252495087/openai-pulls-ai-voice-that-was-compared-to-scarlett-johansson-in-the-movie-her>.

27. Duane C. Pozza, *CES 2025: FTC Commissioners Discuss Approach to AI in New Administration*, Wiley (Jan. 13, 2025), <https://tinyurl.com/4s76a4v8>.

28. *In re Rytr LLC*, Docket No. C-4806 (FTC Dec. 16, 2024); Press Release, Fed. Trade Comm’n, *FTC Approves Final Order Against Rytr, Seller of an AI “Testimonial & Review” Service, for Providing Subscribers with Means to Generate False and Deceptive Reviews* (Dec. 18, 2024), <https://www.ftc.gov/news-events/news/press-releases/2024/12/ftc-approves-final-order-against-rytr-seller-ai-testimonial-review-service-providing-subscribers>.

29. Dissenting Statement of Comm’r Andrew N. Ferguson at 1, *In re Rytr LLC*, Matter No. 2323052 (Sept. 25, 2024).

30. Kyle Wiggers, *President Trump Repeals Biden’s AI Executive Order*, TECHCRUNCH (Jan. 20, 2025), <https://tinyurl.com/3br9ut8k>.

31. AM. BAR ASS’N, *RESOLUTIONS WITH REPORTS TO THE HOUSE OF DELEGATES, 2025 MIDYEAR MTG.* (Feb. 3, 2025), https://www.americanbar.org/content/dam/aba/administrative/house_of_delegates/ebook-of-resolutions-with-reports/2025-midyear-resolution-ebook.pdf.

From the Science & Technology Law Section



Artificial Intelligence: Legal Issues, Policy, and Practical Strategies

Edited by Cynthia H Cwik, Christopher A Suarez, and Lucy L Thomson

The book covers a wide range of important topics concerning AI and the law, and provides practical advice to attorneys on how to navigate these complex and rapidly evolving issues.

The topics include:

- A primer on key AI concepts
- Overarching legal issues, including AI governance and ethics
- AI and law practice management, including the use of digital assistants, AI use cases for lawyers, and paths to the responsible use of AI by lawyers
- AI and the courts
- Possibilities of AI as a tool to close the justice gap
- Intellectual property issues, as well as issues at the intersection of AI and the Internet of Things (IoT)
- The impact of AI on employment law and the workplace
- Specific types of AI risks, including cybersecurity, privacy, and national security
- AI and the law from the perspective of practitioners in Europe
- The long-term future of AI

ISBN: 978163905493

Product Code: 5450086

July 2024, 358 Pages, 7x10
Paperback/eBook

Non-Member: \$149.95

ABA Member: \$134.95

Section Member: \$119.95



2 EASY WAYS TO ORDER

ShopABA.org OR (800) 285-2221