



INSIGHTS

Artificial Intelligence

The Generative Slate: Two Courts Find Fair Use in GenAI Training

Training generative AI with unlicensed works deemed transformative in two cases, but for different reasons and with different implications

By Farrah Vazquez, Adrian F. Vallens,
Nicolas A. Jampol, Jonathan Segal, and
Brad Miller

06.27.25

This article is part of DWT's The Generative Slate series. It explores the use of generative AI in the production and distribution of content.

After nearly two years since the first lawsuit involving generative AI (GenAI) came on the scene—and dozens of lawsuits later—we're now receiving highly anticipated decisions in the battle over whether training GenAI models on unlicensed materials constitutes fair use. In the first fair use decisions of their kind, two courts in the Northern District of California (a battleground for the majority of the GenAI litigation) found that training a GenAI model on copyrighted books was a fair use, but for different reasons, in different circumstances. We discuss these decisions and the legal implications for users of GenAI below.

Fair Use as a Defense to Copyright Infringement

Fair use is a total defense to copyright infringement. The defense focuses on factors including the nature and purpose of the alleged infringement, including whether the use is "transformative" of the copyrighted work's original purpose. The analysis also takes into consideration the market harm caused by the copying of the work.

A few major fair use decisions in years past helped shed light on how judges might rule in cases involving GenAI training. These include, for example, the 2015 *Google Books* case from the 2nd Circuit. In that case, the court held that Google's digital copying of millions of books to create a text-searchable database—from which readers can access snippets, but not the entire books—was transformative because it "augments public knowledge by making available information *about* Plaintiffs' books without providing the public with a substantial substitute for" the originals. *Authors Guild v. Google, Inc.*, 804 F.3d 202, 207-208 (2nd Cir. 2015).

Ten years later, a Delaware federal district court went the opposite direction, declining to find fair use in the 2025 [*Ross Intelligence*](#) case involving an AI tool (but not a *generative* AI tool). Ross Intelligence built a

platform that retrieves relevant legal decisions in response to questions from users. The platform trained on proprietary features of Westlaw's legal research service called Headnotes and Key Numbers, which summarize and organize key points of law and case holdings. The court decided that copying Westlaw's materials was *not* transformative, finding that Ross Intelligence was merely looking to supplant Westlaw's entire *raison d'être* and even rival Westlaw. *Thomson Reuters Ent. Centre GMBH v. Ross Intelligence Inc.*, 765 F. Supp. 3d 382, 397-98 (D. Del. 2025).

Fair Use in GenAI Litigation

Now, some defendants in the current slate of generative AI litigation are arguing that the use of unlicensed copyrighted material in GenAI platforms serves a distinct purpose from that of the works owned by plaintiffs, primarily consisting of books, visual art, news articles, song lyrics, and other expressive works. As these cases progress, defendant AI developers are likely going to be relying on the *Google Books* case, while distinguishing themselves from the specifics of the *Ross Intelligence* case. Plaintiffs, on the other hand, claim that these generative AI tools are competing with their services and were built by copying their copyright-protected works without permission or payment.

Courts Find Training GenAI Models Is a Fair Use

This week brought the very first fair use decisions in the GenAI litigation—both in cases involving training on copyrighted books.

In *Bartz et al. v. Anthropic*, Case No. 3:24-cv-05417 (N.D. Cal.), the court granted Anthropic's motion for summary judgment on the issue of fair use regarding its model's training, finding that Anthropic's use of copyrighted books for training was "exceedingly transformative."

By way of background, Anthropic is the developer of Claude, a family of large language models (LLMs) that generate text in response to user prompts. A group of authors sued Anthropic for ingesting their copyrighted books as part of the LLM's training and for maintaining a library of copies of the authors' books. Anthropic moved for summary judgment, asking the court to dismiss the claims because its activity qualified as a fair use.

The court agreed with Anthropic on the training aspect of the case. It held that using books to train Claude was transformative—and "spectacularly so"—because Anthropic simply used the works "to iteratively map statistical relationships between" the text fragments in the books so that Claude could "return *new* text outputs as if it were a human reading prompts and writing responses" rather than provide users with excerpts from the books themselves. The court drew an analogy to the 2015 *Google Books* case, finding that Anthropic put in place guardrails to ensure that no infringing content ever reached users, similar to how Google had implemented guardrails to ensure limits on how much of a book's text any one user could view in its database. In that way, the court found that training on the books served a different purpose than the books themselves. On the other hand, the court found that Anthropic's practice of maintaining what the court described as a central library for pirated books for undefined purposes was not a fair use. The claim pertaining to the central library will proceed to trial.

Bartz wasn't the only fair use decision this week. Another California federal district court in *Kadrey et al. v. Meta*, Case No. 23-cv-03417 (N.D. Cal.) found that training Meta's model, LLaMA, on copyrighted books was a fair use, albeit for a different reason. Here, the court agreed that training was transformative, but that the case involved "the creation of a product with the ability to severely harm the market for the works being

copied," which the court stated was the most important fair use factor. The court then found that the authors in the case presented no evidence regarding market dilution, thus the factor cut in Meta's favor.

The decisions in *Bartz* and *Kadrey* are just the first of potentially dozens to come, each with their own unique facts and circumstances. For example, some involve allegations that AI model *outputs* contained snippets from the copyrighted works that it trained on (a fact the court in *Bartz* suggested could influence the outcome). Some may involve evidence of harm to the market for the copyrighted works (which the court in *Kadrey* suggested could make a significant difference). Others do not involve companies that maintain copies of the allegedly infringed works.

What Does This Mean for Content Producers and Distributors?

Companies that produce or distribute content are becoming more involved in AI copyright cases. In recent weeks, two major studios sued AI developer Midjourney for copyright infringement. Other studios have agreed to license their content to AI developers to train on their works. The [U.S. Copyright Office recently weighed in](#) on the hotly debated fair use question, concluding only that "some uses of copyrighted works for generative AI training will qualify as fair use, and some will not."

While the *Bartz* and *Kadrey* decisions provide a hint at what's to come, they demonstrate that different courts may focus on different factors, and that different circumstances in individual cases may drive the outcome. Some courts may focus on the transformativeness of the training, others on the alleged market harm, and still others on whether the AI-generated outputs are substantially similar to any existing

copyright-protected work. These cases also confirm there won't be consistent judicial guidance on these issues anytime soon.

As a result of this ongoing uncertainty and rapidly changing landscape, producers and distributors should monitor these cases to determine legal risk and continuously evaluate their use of GenAI on a case-by-case basis to guard against potentially infringing outputs, while GenAI developers will need to assess their practices for copying, training, and quality control, and some may need to consider deploying additional guardrails.

For any questions about the above, please contact the authors. Otherwise, stay tuned for the next update from The Generative Slate.

Related Insights

04.23.25
INSIGHTS

Utah Enacts Multiple Laws Amending and Expanding the State's Regulation of the Deployment and Use of Artificial Intelligence

02.14.25
INSIGHTS

Thomson Reuters v. Ross Intelligence: Copyright, Fair Use, and AI (Round One)

10.02.24

INSIGHTS

California Governor Signs AI Governance Measures into Law—But Vetoes Controversial AI Safety Bill