Landmark Ruling on AI Copyright: Fair Use vs. Infringement in *Bartz v. Anthropic*

JUNE 30, 2025

MATTHEW L. FINKELSTEIN*, DANIELLE W. BULGER, DAN JASNOW, NATASHA WEIS

Share This Page   EMAIL   LINKEDIN   X   FACEBOOK

In one of the first substantive decisions analyzing whether the use of copyrighted works to train large language models (LLMs) for generative artificial intelligence (AI) services is infringing or a fair use, Judge William Alsup issued a split decision in his summary judgment order. *See Bartz et al. v. Anthropic PBC, No. 3:24-cv-05417* (N.D. Cal. Aug 19, 2024).

Listen to this article now
Powered by **Trinity Audio**

A文   🔁10   🔁10   1.0x

00:00   07:44

Judge Alsup ruled that (1) Anthropic's use of the books at issue to train LLMs for the purpose of returning new text outputs is "spectacularly" transformative and therefore a fair use, (2) Anthropic's digitization of books it purchased in print form for use as part of its central library was a fair use because the digital copies were a replacement of the print copies it discarded after digitization, and (3) Anthropic's use of "pirated" copies of books in its central library was infringing. The order came on Anthropic's early motion for summary judgment on the question of fair use regarding certain of Anthropic's uses of the books at issue. Numerous other issues remain for trial and are not discussed in this alert.

Background

Anthropic used millions of copyrighted books to train its Claude LLMs for use with its AI services capable of generating writings that mimic the writing style of humans. In preparing to train its LLMs for research and product development, Anthropic compiled a "central library" of "all the books in the world" to retain "forever." Novels and non-fiction titles written by the author-plaintiffs in the case and owned by the author-plaintiffs or their companies were among the sets of books and text in the central library.

Anthropic sourced content for its library in various ways. It downloaded for free millions of pirated copies of books in digital form. In addition, it purchased millions of copyrighted books (some overlapping with those acquired from pirate sites), removed the bindings, scanned and stored the works in a digitized searchable format, and then discarded the paper originals.

Each work selected for training was copied in four main ways and, as Anthropic admitted, so many times that it would be impractical to estimate. However, according to the facts in the record, the training copies were not disseminated to the outside world. Rather, when each LLM was put into a public-facing version of Claude, it was combined with other software that filtered user inputs to the LLM and filtered outputs from the LLM back to the user. Ultimately, the plaintiffs did not allege any infringing copy of their works was or would ever be provided to users by the Claude service. The plaintiffs did, however, claim Anthropic infringed their copyrights by (1) pirating copies of their works for Anthropic's library and (2) reproducing their works to train Anthropic's LLMs.

In support of their copyright infringement claims, the plaintiffs alleged, among other arguments, that use of their books to train Anthropic's LLMs could result in the production of works that compete and displace demand for their books. In addition, the plaintiffs alleged Anthropic's unauthorized use has the potential to displace an emerging market for licensing the plaintiffs' works for the purpose of training LLMs.

The Court's Fair Use Analysis

In its fair use analysis, the court differentiated between Anthropic's copying of millions of copyrighted materials for the purpose of training its LLMs, and Anthropic's retention of copies of books and text for building its central library. The court concluded that use of the books at issue to train Anthropic's LLMs was "exceedingly transformative" and a fair use under Section 107 of the Copyright Act. Specifically, the court noted that authors cannot exclude others from using their works to learn. It noted that, for centuries, people have read and re-read books, and that the training was for the purpose of creating something different, not to supplant the works.

With respect to the digitization of the books purchased in print form by Anthropic, the Court concluded this was fair use. It reasoned that, because the new digital copies were not redistributed, but rather, simply, convenient space-saving replacements of the discarded print copies, the digitization resulted in a "format change" that did not relate to one of the exclusive rights granted under the Copyright Act and reserved to authors to exploit. Anthropic purchased the print copies "fair and square[,]" thus, the digital copy should be treated "just as if the purchased print copy had been placed in the central library[,]" according to the Court.

The court, however, reached a different conclusion with respect to the pirated copies. Because Anthropic never paid for the pirated copies, the court thought it was clear the pirated copies displaced demand for the authors' works, copy for copy. The fact that pirated copies would later be used for a purpose the court found to be transformative — training LLMs — did not dissuade them against finding no fair use. That finding applied equally to the books Anthropic initially pirated and later purchased in print form. Per the court's opinion, "no damages from pirating copies could be undone by later paying for copies of the same works."

In response to the authors' contention that training LLMs displaced (or will) an emerging market for licensing their works for the narrow purpose of training LLMs, the court's view was that such a market is not one the Copyright Act entitles the authors to exploit.

Among matters left open for determination at trial is whether the making of copies of the books from Anthropic's central library copies — for use other than training the LLMs — is infringing or a fair use, and resulting damages.

Because the authors never alleged the outputs infringed their rights in this case, the court focused solely on the training data, or "inputs." However, in several instances the court emphasized that the case would be significantly different if the outputs created by Anthropic's LLMs were infringing, seemingly hinting to a number of the other pending copyright infringement cases involving AI outputs (for example, the **Midjourney case**, among others.).

Conclusion

Although this decision serves as one of the first substantive court orders addressing the question of fair use in the context of using copyrighted works in datasets to train AI, it is not the last. Days later, in the District Court for the Northern District of California, Judge Vince Chhabria **granted** another AI developer's motion for summary judgment on the issue of fair use related to training LLMs. But, in doing so, applied a different fair use analysis than Judge Alsup. *Kadrey v. Meta Platforms, Inc.*, No. 23-CV-03417-VC, 2025 WL 1752484 (N.D. Cal. June 25, 2025). That ruling foreclosed a group of authors' copyright infringement claims against an AI developer, who the plaintiffs claimed downloaded their copyrighted books from "shadow libraries" and used their works to train LLMs.

On balance, these recent decisions demonstrate the complexities and fact-intensive nature of the fair use analysis. They are certain to be closely scrutinized by AI developers, rights' holders, and other courts. Both summary judgment decisions are dispositive only with respect to the specific claims of the plaintiffs in these cases. It remains to be seen how courts will decide in cases where plaintiffs allege copyright infringement resulting from an AI service generating output that is substantially similar to works protected by copyright law.

Contacts

Matthew L. Finkelstein*

PARTNER

Danielle W. Bulger

PARTNER

Dan Jasnow

PARTNER AND AI GROUP CO-LEADER

Natasha Weis

ASSOCIATE

Related Industries

AI, Metaverse & Blockchain

Media & Entertainment

Related Practices

Copyright

Continue Reading

BIS Issues Two General Authorizations and More FAQs on the Connected Vehicles Rule

JUNE 30, 2025 | SYLVIA G. COSTELLOE, MAYA S. COHEN

FDA Unveils a Tool for Prioritizing Food Chemicals for Post-Market Assessment

JUNE 30, 2025 | ROBERT G. EDWARDS, PH.D.

Restructuring DC Health Care Alliance With New Eligibility Rules and Timelines for 2025

JUNE 30, 2025 | DAVID P. GROSSO, JON S. BOUKER

All Insights from Alerts