

Client Alert - Fair Use, Copyright, and AI Training: Key Insights from the Anthropic Decision

On June 23, 2025, Judge William Alsup of the U.S. District Court for the Northern District of California issued a significant opinion in *Bartz et al. v. Anthropic PBC*, addressing whether using copyrighted works to train large language models (LLMs) constitutes fair use. This decision is one of the most detailed judicial treatments to date on how copyright law applies to AI training and data acquisition.

Background and Nature of the Ruling

This ruling came in response to a summary judgment motion filed by Anthropic PBC, which sought a court determination — prior to class certification — that its uses of copyrighted books were protected under the fair use doctrine in Section 107 of the Copyright Act.

- Procedural Posture: Judge Alsup granted Anthropic leave to file a pre-class certification motion focused solely on fair use, enabling an early resolution of foundational legal questions in the case.
- Scope and Burden: The court evaluated whether specific categories of Anthropic's conduct — including training LLMs on copyrighted materials, digitizing purchased books, and building an internal library of pirated copies — qualified as fair use. As the movant, Anthropic bore the burden of demonstrating fair use on undisputed facts, with all ambiguities resolved in favor of the plaintiffs.
- Outcome:
 - Summary judgment granted for Anthropic with respect to:
 - Using copyrighted works for training LLMs.
 - Digitizing lawfully purchased print books for internal use.
 - Summary judgment denied as to:
 - The acquisition and retention of pirated books, which the court found could not be justified under fair use, should proceed to trial.

This decision is not a final judgment in the case, but it narrows the issues for trial and sets out important boundaries for how courts may analyze fair use in the context of AI.

Fair Use Applies to Training LLMs with Copyrighted Works

The court held that using copyrighted books to train LLMs was a fair use, emphasizing:

- Transformative Use: The use was “spectacularly transformative.” The LLMs did not reproduce or distribute the texts, but instead used them to learn language patterns, much like a person reading literature to become a better writer.
- No Public Distribution: Plaintiffs did not allege that any of their works were reproduced or leaked to users through the AI system, and the court credited Anthropic's implementation of output filters to prevent such behavior.

- No Market Substitution: The LLMs did not serve as a replacement for the books in the market, nor did their training encroach on the expressive purpose of the original works.

Format Conversion of Purchased Books Is Also Fair Use

Anthropic purchased millions of print books, scanned them into searchable PDFs, and discarded the originals. The court treated this as a separate use, and also fair:

- Lawful Acquisition: Anthropic bought the books legitimately.
- No Redistribution: The digital copies were stored internally and not shared.
- Functional Transformation: The digitization was for space-saving and internal searchability — non-expressive and non-commercial purposes.

Judge Alsup distinguished this use from more substantive, transformative uses, such as LLM training. While not transformative in the expressive sense, the format shift was permitted under fair use because it merely substituted one lawful copy (in print) for another (in digital format), with no surplus or public exposure.

Use of Pirated Copies to Build a Central Library Is Not Fair Use

In contrast, the court rejected fair use for Anthropic's early conduct in downloading over seven million pirated books from sites like Books3, LibGen, and PiLiMi to populate its central "research library":

- Acquisition Matters: Even if the ultimate use (training) could be transformative, the initial act of piracy could not be excused.
- Retained Without Use: Anthropic retained many pirated works indefinitely, even those not used in model training.
- Lacked Separate Justification: The central library was itself an infringing use — a general-purpose archive of unlawfully acquired content with no independent fair use rationale.

A trial will now proceed to determine liability and damages related to these pirated materials.

Authors Cannot Claim Control Over Every Emerging Licensing Market

Plaintiffs argued that training AI on copyrighted works displaced a potential new licensing market for such uses. The court acknowledged the possibility of such a market emerging but decisively held that it was not one the Copyright Act guarantees to authors.

- Legal Principle: Courts have long held that copyright holders cannot monopolize all transformative or orthogonal uses of their works. See, e.g., *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605 (2d Cir. 2006); *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).
- Licensing Expectations: The mere fact that authors wish to license a use — such as for AI training — does not make it a market they are legally entitled to control.

"Such a market for that use is not one the Copyright Act entitles Authors to exploit." — Judge Alsup

Key Takeaways for Clients Across Industries

- For AI Developers: Focus on lawful acquisition of training data and use filtering to prevent output-based infringement. Do not rely on fair use to excuse shortcuts in content acquisition, especially when it comes to piracy.
- For Copyright Owners: While training on your works may feel exploitative, courts currently draw sharp distinctions between public-facing uses and internal, transformative uses. Enforcement efforts should focus on acquisition and output. While the court rejected the plaintiffs' licensing argument in connection with fair use factor four, the sheer amount of "legal" copying necessary to build a library like Anthropic's suggests that a licensing model remains a viable means of revenue generation.
- For All Organizations: Review your internal handling of copyrighted material, especially scanned books, third-party datasets, or content used in machine learning workflows.

What's Next

A trial will proceed to determine:

- Damages (actual or statutory) for the pirated copies.
- Whether infringement was willful.
- Whether other uses of the central library also constitute infringement.

The ruling reinforces that each use of a copyrighted work must be evaluated separately under fair use, including intermediate copies and back-end retention.

A copy of the order is available at [Judge-Alsup-order-on-fair-use-and-infringement-Jun-23-2025.pdf](#)

If you would like assistance in evaluating your organization's copyright exposure, AI model training practices, or content licensing policies in light of this decision, please contact:

Lawrence R. Robins, Partner, Chair Brand Management Practice Group, at larry.robins@fisherbroyles.com.

About FisherBroyles, LLP

Founded in 2002, FisherBroyles, LLP is the first and one of the world's largest distributed law firm partnerships. The Next Generation Law Firm® has grown to hundreds of partners practicing in 29 markets globally. The FisherBroyles' efficient and cost-effective Law Firm 2.0® model leverages talent and technology instead of unnecessary overhead that does not add value to our clients, all without sacrificing BigLaw quality. Visit our website at www.fisherbroyles.com to learn more about our firm's unique approach and how we can best meet your legal needs.

These materials have been prepared for informational purposes only, do not constitute legal advice, and under applicable rules of professional conduct governing attorneys in various jurisdictions, may be considered advertising materials. This information is not intended to and does not create an attorney-client or similar relationship. Whether you need legal services and which lawyer you select are important decisions that should not be based on these materials and information alone.

© 2025 FisherBroyles, LLP