

AI Wins Big on "Fair Use," But Judge Slams Brakes on Piracy in Landmark Anthropic Copyright Ruling

Article By:

Andrew R. Lee

Jason M. Loring

Graham H. Ryan

Timothy P. Scanlan, Jr.

A federal judge has handed the AI industry a massive victory. Still, it came with a crucial catch: innovation can't be built on a foundation of theft, and AI systems must earn their authority through legitimate means.

In a closely watched case, a US District Judge [ruled](#) that AI company Anthropic's use of copyrighted books to train its powerful AI model, Claude, was "exceedingly transformative" and qualified as a legal "fair use." The decision is a game-changer for AI developers who argue that learning from vast datasets is essential for innovation.

However, the judge drew a sharp line in the sand, ruling that Anthropic's separate act of downloading and storing millions of books from "pirate sites" was *not* a fair use and that the company will have to face a trial for it.

This nuanced decision strikes a balance in the high-stakes battle between copyright holders and the rapidly evolving world of artificial intelligence, reflecting a growing recognition that *how* AI systems acquire their capabilities matters as much as *what* they can do with them.

When AI models shape human understanding and decision-making at an unprecedented scale, the legitimacy of their knowledge sources becomes a question of technological integrity, not just legal compliance.

The Core of the Case: A Tale of Three Authors

A trio of authors brought the lawsuit:

- **Andrea Bartz**, author of thrillers like *The Lost Night* and *We Were Never Here*

-
- **Charles Graeber**, who penned the true-crime story *The Good Nurse* and the medical chronicle *The Breakthrough*
 - **Kirk Wallace Johnson**, author of nonfiction works including *The Feather Thief* and *The Fishermen and the Dragon*

The authors alleged that Anthropic, a multi-billion-dollar frontier large language model (LLM) backed by Amazon and Google, built its AI by infringing on their copyrights by feeding their books into Claude without permission or payment. Section 107 of the Copyright Act identifies four factors for determining whether a given use of a copyrighted work is a fair use: (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.

The Ruling: A "Quintessentially Transformative" Use

Judge William Alsup of the Northern District of California sided with Anthropic on the most critical question — the use of copyrighted works to train an AI model. He reasoned that Anthropic's approach to the books was not to replace them, but to learn from them to create something entirely new.

In a powerful analogy, Judge Alsup wrote that the process was **"quintessentially transformative. Like any reader aspiring to be a writer, Anthropic's LLMs trained upon works not to race ahead and replicate or supplant them — but to turn a hard corner and create something different."**

The court emphasized that the AI did not simply spit out copies of the authors' work. The judge noted that Anthropic's models **"have not reproduced to the public a given work's creative elements, nor even one author's identifiable expressive style."** Because the final product didn't compete with or replace the original books, the training process was deemed fair.

This part of the ruling is a significant relief for AI companies, who have long argued their training methods are a modern form of research and learning.

The Catch: "You Can't Just Bless Yourself"

While the ruling on AI *training* was a clear win for Anthropic, the judge took a much dimmer view of how the company acquired a large portion of its data. The court found that before Anthropic began purchasing and scanning millions of physical books, it first downloaded over seven million books from known pirate libraries, such as LibGen.

Anthropic kept these books in a massive "central library" to use for "research," retaining them "forever" even if they were never used for training. The court's reasoning suggests that AI systems can legitimately "read" human cultural output to develop their capabilities, but only when that reading occurs through recognized channels of access and permission.

Judge Alsup rejected the idea that a future fair use can excuse initial theft. He quoted Anthropic's own lawyer's argument back at them in his decision: **"You can't just bless yourself by saying I have a research purpose and, therefore, go and take any textbook you want. That would destroy the academic publishing market if that were the case."**

The ruling implicitly acknowledges that AI systems do not operate in isolation — they function as intermediaries that shape how humans access and understand information. When these systems are built on illegitimately acquired content, they potentially perpetuate unauthorized appropriation at scale, influencing human choices based on improperly obtained knowledge.

The ruling was blunt about the piracy:

"This order doubts that any accused infringer could ever meet its burden of explaining why downloading source copies from pirate sites that it could have purchased or otherwise accessed lawfully was itself reasonably necessary to any subsequent fair use."

For Anthropic, this means it will face a trial for damages. The judge scheduled a trial to determine the damages Anthropic may have to pay for the infringement. He concluded that a later purchase doesn't erase the initial crime: **"That Anthropic later bought a copy of a book it earlier stole off the internet will not absolve it of liability for the theft, but it may affect the extent of statutory damages."**

What This Means for AI and Authors

This landmark decision offers both sides a partial victory and sets a critical precedent.

- **For AI Companies:** The ruling validates the core argument that training LLMs on copyrighted material can be a transformative fair use. This suggests a legal green light for the fundamental process that powers generative AI, provided the material is lawfully acquired. This creates a framework where AI systems can legitimately learn from human cultural expression while respecting the rights of creators. It also establishes, however, that the means of acquisition matter — AI systems that will increasingly mediate human access to information must themselves be developed through legitimate channels.
- **For Authors and Creators:** The court proclaims that AI companies are not above the law. They cannot simply scrape content from pirate sites to build their models. This creates an incentive for AI developers to license content or find other legitimate ways to source their training data.

Looking Forward: Beyond Legal Compliance

This ruling arrives as AI systems become more sophisticated mediators of human knowledge and decision-making. The court's emphasis on legitimate acquisition suggests a recognition that AI development practices have implications that extend beyond copyright law. When AI systems can influence human understanding through their responses, the integrity of their training processes becomes a matter of technological accountability.

The decision may also influence how courts approach other aspects of AI development, where the methods used to create AI capabilities affect their legitimacy as information intermediaries. As AI systems become more integrated into research, education, and professional decision-making, questions about the integrity of their development processes will likely extend beyond copyright to encompass broader concerns about technological transparency and accountability.

If upheld, the decision will stand for the notion that AI can read the world's books to learn, but it first

needs a library card. More broadly, it suggests that as AI systems become powerful mediators of human knowledge and choice, their authority must be earned through legitimate means. The future of AI will not be built on piracy — either of content or of the trust necessary for AI systems to serve as reliable partners in human decision-making.

Over time, Anthropic came to value most highly for its data mixes books like the ones Authors had written, and it valued them because of the creative expressions they contained. Claude's customers wanted Claude to write as accurately and as compellingly as Authors. So, it was best to train the LLMs underlying Claude on works just like the ones Authors had written, with well-curated facts, well-organized analyses, and captivating fictional narratives — above all with “good writing” of the kind “an editor would approve of.” Opinion, p. 6

© 2025 Jones Walker LLP

National Law Review, Volume XV, Number 175

Source URL: <https://natlawreview.com/article/ai-wins-big-fair-use-judge-slams-brakes-piracy-landmark-anthropic-copyright-ruling>