Federal Courts Find Fair Use in AI Training: Key Takeaways from Kadrey v. Meta and Bartz v. Anthropic

July 11, 2025 | Insights

By [Arthur Gollwitzer](#)

In recent days, two federal judges in the Northern District of California issued significant decisions covering the intersection of artificial intelligence (AI) and copyright law. Specifically, in *Bartz v. Anthropic PBC* and *Kadrey v. Meta Platforms, Inc.*, the courts addressed whether the use of copyrighted works to train generative AI models constitutes fair use under Section 107 of the Copyright Act. While both judges ruled in favor of the AI developers based on the "fair use" doctrine, their reasoning and cautions highlight the complexity and evolving nature of this legal frontier.

Bartz v. Anthropic: Transformative Use and the Limits of Fair Use

In *Bartz v. Anthropic*, Judge William Alsup confronted the question of whether Anthropic's use of millions of copyrighted books to train its Claude AI model was protected by the fair use doctrine. Judge Alsup focused heavily on the transformative nature of AI training. He concluded that using books to train a generative AI system is "exceedingly transformative," likening it to how a human might read, internalize, and later draw upon a book's themes and style to create new works. The court emphasized that the AI's outputs did not reproduce or closely mimic the plaintiffs' works, and the training process itself was fundamentally different from the original purpose of the books.

Judge Alsup also addressed the issue of digitizing lawfully purchased print books, concluding that converting these books into digital format for internal research and training purposes constituted fair use. He reasoned that this process simply substituted the physical copy with a more accessible digital version, without generating new works or sharing the material outside the organization.

However, the court made a clear distinction regarding Anthropic's acquisition and retention of pirated books. While the use of these books for training was transformative, the creation and maintenance of a permanent, general-purpose digital library of pirated works was not protected by fair use. Judge Alsup made clear that "pirating copies to build a research library without paying for it, and to retain copies should they prove useful for one thing or another, was its own use—and not a transformative one." Therefore, while the training use was excused, the underlying act of piracy was not.

Kadrey v. Meta: Market Harm as a Decisive Factor

In *Kadrey v. Meta*, Judge Vince Chhabria took a somewhat more cautious approach. While acknowledging the transformative nature of using copyrighted works to train large language models (LLMs) like Meta's Llama, Judge Chhabria emphasized that transformativeness alone does not guarantee fair use. The most significant factor, he explained, is how the use impacts the market value or potential market for the original works.

Judge Chhabria began with the proposition that, in most cases, using copyrighted works to train AI models without permission likely will constitute infringement, especially if it undermines the market for those works. In the case before the court, however, the thirteen plaintiff authors did not provide evidence that Meta's use of their works resulted in any market harm, either by causing the AI to reproduce substantial portions of their books or by negatively affecting the market for licensing books as AI training data. Moreover, the court found that Meta's Llama model could not output more than trivial snippets of the plaintiffs' works and the plaintiffs had not established a cognizable market for licensing books for AI training.

Importantly, Judge Chhabria's ruling was fact-specific. He cautioned that his decision "does not stand for the proposition that Meta's use of copyrighted materials to train its language models is lawful. It stands only for the proposition that these plaintiffs made the wrong arguments and failed to develop a record in support of the right one." The court also noted that Meta's acquisition of the books from "shadow libraries" (unauthorized online repositories) did not, in itself, preclude a finding of fair use for the training purpose, though the manner of acquisition could be relevant in other contexts.

Key Questions Remain for AI Developers and Copyright Holders

While both decisions are positive developments for AI developers, they do not provide blanket endorsements of certain AI industry practices. The courts' analyses underscore that fair use in the context of AI training is highly fact-dependent and will turn on several critical questions:

- **How were the training materials acquired?** Piracy or unauthorized copying for the purpose of building a permanent library may constitute copyright infringement, even if subsequent use for training may constitute a fair use.
- **How many copies were made, and for what purpose?** The distinction between making copies for transformative training versus for general archival purposes is important.
- **What is done with the copies?** Retaining unauthorized copies of copyrighted works for future, unspecified uses weighs against finding a fair use.
- **What kind of output can the AI model produce?** If the AI can regurgitate substantial portions of copyrighted works, or if its outputs serve as market substitutes for the original works, the fair use doctrine is less likely to apply.
- **What is the impact on the market for the original works?** Plaintiffs must show actual or likely market harm, not just speculative or theoretical harm.

Conclusion

The *Bartz* and *Kadrey* decisions provide important, if initial, guidance for the rapidly evolving field of AI and copyright law. They confirm that training generative AI models on copyrighted works can, under certain circumstances, qualify as fair use, especially when the use is transformative and does not harm the market for the original works. The courts, however, emphasized that not every use will be permitted, especially in cases involving piracy or harm to the market.

For AI developers, these rulings highlight the importance of careful, fact-specific legal analysis and the need to consider how training data is acquired, used, and managed. For copyright holders, the decisions highlight the necessity of developing thorough evidence of market harm and understanding the nuances of fair use in the AI context.

The opinions expressed are those of the authors and do not necessarily reflect the views of the firm, its clients, or any of its or their respective affiliates. This article is for informational purposes only and does not constitute legal advice. For questions, please contact a member of the Intellectual Property practice.

Meet Art

Arthur Gollwitzer combines experience gained as a federal prosecutor with nearly 30 years in private practice to represent clients in patent, copyright, trademark, and trade secret disputes, including jury trials and appeals. In addition to defending and enforcing intellectual property rights, Arthur is highly experienced in investigations and criminal litigation. Arthur writes and speaks frequently on patent litigation reform and other

intellectual property issues.