# Meta faces pivotal copyright battle over AI training on pirated books



In a landmark legal proceeding set to unfold in San Francisco on May 1, 2025, Meta Platforms Inc. faces a critical test of the boundaries of copyright law as it pertains to generative artificial intelligence (AI). The U.S. District Court for the Northern District of California will hear arguments in the first of multiple lawsuits challenging the use of copyrighted materials for training AI models, with a ruling poised to shape the future of AI development and copyright enforcement.

The case, Kadrey v. Meta Platforms Inc., brings together a class of authors including high-profile figures such as comedian Sarah Silverman, journalist Ta-Nehisi Coates, and Pulitzer Prize-winning novelist Andrew Sean Greer. These plaintiffs assert that Meta engaged in widespread copyright infringement by using pirated books to train its AI model known as Llama. According to the plaintiffs, Meta knowingly downloaded millions of copyrighted books from so-called "shadow libraries" — notorious online piracy networks like Library Genesis and Z-Library — without proper authorisation.

The legal dispute centres on whether Meta’s alleged practice constitutes direct copyright infringement or qualifies as fair use under U.S. law. Meta denies wrongdoing, contending that its copying of these materials to train the Llama model should be protected as fair use, a legal doctrine that permits limited use of copyrighted works without permission for transformative purposes. The company argues that Llama, as a generative AI, is fundamentally distinct from a book and serves a transformative function, which they say justifies their use of the copyrighted texts in training the AI.

Meta maintains that the origins of the training data should not affect the fair use analysis, regardless of how the books were obtained. In legal briefs submitted to the court, Meta emphasised that “Llama is nothing like a book; it is not meant to be read,” underscoring the transformative and non-consumptive nature of the AI.

The authors counter that evidence presented during discovery — including emails and depositions — indicates Meta abandoned initial licensing arrangements and turned to downloading pirated materials via Torrent, a peer-to-peer file-sharing protocol that also redistributes the files. They argue that Meta’s use of shadow libraries, many of which have been sanctioned by courts for piracy, undermines the company’s fair use defence and demonstrates bad faith.

Legal experts note that the case presents “dramatically different framings of what they consider to be the undisputed facts,” according to Edward Lee, an intellectual property law professor at Santa Clara University. The court’s decision, overseen by Judge Vince Chhabria, may set a significant precedent for how copyright law applies to the rapidly evolving AI industry.

Kevin Madigan, senior vice president of Policy and Government Affairs at the Copyright Alliance, highlighted the broader implications, saying the ruling “could send ripples throughout other cases” pending against AI firms. These include lawsuits against OpenAI, Anthropic, and Google, as well as litigation by music publishers and visual artists over the use of copyrighted works in AI training.

The AI industry has pointed to past cases such as Authors Guild v. Google Inc., where Google's scanning of books to create a searchable database was deemed fair use, as legal precedent supporting transformative uses. However, rights holders distinguish these cases from generative AI technology, arguing the latter’s functionality and scale present new legal challenges.

Meta is represented by Cooley LLP, Cleary Gottlieb Steen & Hamilton LLP, and Paul, Weiss, Rifkind, Wharton & Garrison LLP, while the plaintiffs are represented by a coalition of firms including Boies Schiller Flexner LLP and Joseph Saveri Law Firm LLP.

The case is closely watched as it may determine whether AI companies can continue their reliance on unlicensed copyrighted content for training without incurring legal liability. A ruling against Meta could expose the company to potentially billions of dollars in damages and influence ongoing and future litigation against other AI developers.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://www.reuters.com/legal/litigation/meta-says-copying-books-was-fair-use-authors-ai-lawsuit-2025-03-25/> - Meta argues that its use of authors' books to train the Llama AI model falls under the 'fair use' doctrine, a central point in the Kadrey v. Meta lawsuit.
2. <https://www.casemine.com/judgement/us/678c8001ad6b3e5bf58a732f> - The court's judgment in Kadrey v. Meta Platforms, Inc. addresses the plaintiffs' claims of direct and vicarious copyright infringement, as well as violations of the Digital Millennium Copyright Act (DMCA).
3. <https://www.ccn.com/news/technology/meta-vs-kadrey-ai-copyright-law/> - The article discusses the progression of the Kadrey v. Meta lawsuit, highlighting the court's dismissal of certain claims and the ongoing legal battle over AI training data.
4. <https://www.loeb.com/en/insights/publications/2023/12/richard-kadrey-v-meta-platforms-inc> - This publication provides an overview of the Kadrey v. Meta case, detailing the authors' allegations and the court's dismissal of most claims except for direct copyright infringement.
5. <https://www.theatlantic.com/technology/archive/2025/03/libgen-meta-openai/682093/?utm_source=apple_news> - The article explores the use of pirated books from platforms like Library Genesis in training AI models, a practice central to the Kadrey v. Meta lawsuit.
6. <https://www.fbm.com/eugene-y-mar/publications/is-the-copyright-threat-to-generative-ai-overhyped-implications-of-kadrey-v-meta/> - This analysis examines the implications of the Kadrey v. Meta decision on the intersection of AI language model training and copyright law.
7. <https://news.google.com/rss/articles/CBMingFBVV95cUxONnJoQ2tMN2lRR1FEVWRIajV3TWFqZ2RWNEVxSjI4NDhnTGQ5OURIZ0hXUzlwakNDMEMxVkwwdHl5SkFSSjFaQUhhZzBNRWRiSnRLNVBkQkUtOG4yeFBkVnRzMGd1QVBKcXF4ekwxMlJZeUdwSFpkQUxJOEJVMFFZeHhlQXJfUDNraDNEZG5wbWdFNTUwTkp4dnNIMkVXUQ?oc=5&hl=en-US&gl=US&ceid=US:en> - Please view link - unable to able to access data