



ALERT · JUNE 27, 2025

Northern District of California Judge Rules That Meta's Training of AI Models is Fair Use

BY Robert D. Carroll Stefan Mentzer Timothy Keegan Ji Young Ahn

Just two days after Judge Alsup issued his fair use decision in *Bartz v. Anthropic*, Judge Chhabria of the Northern District of California granted summary judgment for Meta Platforms in an AI copyright infringement suit. Judge Chhabria ruled that, on the record before him, it was fair use for Meta to train its large language models with copyrighted books.¹ In dicta, however, Judge Chhabria wrote that, had the plaintiffs argued and presented evidence on a “market dilution” theory, they very well might have won, overcoming the transformative purpose and character of Meta’s use and weighing the fair use analysis in favor of the authors.

Judge Chhabria’s decision articulates a novel argument that likely will be litigated in a range of AI fair use cases: using copyrighted works to train LLMs can *both* be highly transformative *and* harm the market for a copyright owner’s work, and thus weigh the analysis squarely against fair use and in favor of infringement. “No matter how transformative LLM training may be, it’s hard to imagine that it can be fair use to use copyrighted books to develop a tool to make billions or trillions of dollars while enabling the creation of a potentially endless stream of competing works that could significantly harm the market for those books.”

Background

Meta develops and offers its Llama generative AI models. To train its LLMs, Meta copied large numbers of copyrighted materials, including books. Meta sourced them from “shadow libraries” – online repositories of copyrighted works available for free download. Meta then used the materials it obtained to train the Llama models, and also to “post-train” the models to keep them from “memorizing” and outputting certain text from training data. Meta did not obtain licenses from the copyright owners to obtain or use these materials.

Thirteen published authors of novels, plays, short stories, memoirs, essays, and nonfiction books sued for copyright infringement. The parties filed cross-motions for partial summary judgment on the issue of fair use. Judge Chhabria issued his decision on June 25, 2025.

Summary judgment for Meta on training AI models

The court weighed the fair use factors in favor of Meta on the question of using copyrighted works to train LLMs.

- **Factor 1 – the purpose and character of the use – for Meta.** The court found Meta’s use “highly transformative.” The purpose of Meta’s copying was to train LLMs, in contrast to the purpose of the plaintiffs’ books, which “is to be read for

entertainment or education.” Meta’s commercial purpose, and its downloads of unauthorized copies from “shadow libraries,” did not tip the factor in favor of the authors. The factor weighed for fair use.

- **Factor 2 – nature of the copyrighted works – for the authors.** The court found the plaintiffs’ works – mostly novels, memoirs, and plays – “highly expressive.” The court rejected Meta’s argument that it used the books to gain access to their “functional elements,” not to capitalize on their creative expression. The factor weighed against fair use.
- **Factor 3 – amount and substantiality of the portion used – for Meta.** The court found the amount of Meta’s copying reasonable in light of its transformative purpose. The factor weighed for fair use.
- **Factor 4 – effect on the marketplace of the copyrighted works – for Meta, but with an important “market dilution” caveat.** The court found in Meta’s favor based on the record before it. It rejected the authors’ arguments that Llama’s outputs would regurgitate their works and substitute for them. It also rejected the argument that Meta harms the authors’ market for licensing AI training. But the court found “far more promising” a third argument – that an LLM can generate works that are similar enough in subject matter in genre that they will compete with the originals and indirectly substitute for them. This “market dilution” argument was compelling for Judge Chhabria in light of Supreme Court precedent that market substitution is the “only harm” that matters under factor 4. Using copyrighted books to train LLMs “might harm the market for those works . . . by helping to enable the rapid generation of countless works that compete with the originals, even if those works aren’t themselves infringing.” But Judge Chhabria recognized the fact-bound nature of this inquiry, and he ultimately found the plaintiff’s presentation “so weak that it does not move the needle, or even raise a dispute of fact sufficient to defeat summary judgment.” For that reason, the court weighed the factor for fair use.

Takeaways

- Meta won on the important point that copying to train LLMs is transformative under factor 1. Paired with Judge Alsup’s decision in *Bartz v. Anthropic* that training LLMs is transformative, the decision suggests that other courts will rule the same way. Factor 1 transformativeness is crucial to the fair use determination.
- But the decision raises a potential new challenge for the fair use defense in AI copyright cases. Judge Chhabria suggested that a “market dilution” analysis under factor 4 is the most critical question and can lead to a finding of no fair use, even where the purpose and character under factor 1 is highly transformative. On that basis, Judge Chhabria described his decision as “limited”: “this ruling 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 decision invites copyright owners to develop the “market dilution” argument in other AI cases.
- The decision also highlights two different approaches to assess fair use for AI. Judge Alsup in *Bartz v. Anthropic* conducted three distinct fair use analyses, looking at (1) training LLMs, (2) converting in-print books to digital copies, and (3) using “pirated” copies to build a “central library.” In contrast, Judge Chhabria performed a single analysis, considering only the end-goal of the copying which was to train LLMs. He declined to separately consider whether it was fair use to download the books or to download them but not use them for LLM training, finding that all of Meta’s downloads “had the ultimate purpose of LLM training.” Other courts considering fair use will have to decide which uses they will consider when assessing questions of fair use.

^[1] *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC (N.D. Cal. June 25, 2025).

CONTACTS

Robert D. Carroll

Partner

rcarroll@goodwinlaw.com

Boston | +1 617 570 1753

Stefan Mentzer

Partner

smentzer@goodwinlaw.com

New York | +1 212 459 7088

Timothy Keegan

Associate

tkeegan@goodwinlaw.com

New York | +1 646 296 1411

Ji Young Ahn

Associate

jahn@goodwinlaw.com

New York | +1 212 813 8832



GOODWIN