Judges Take Divergent Paths to Find AI Training "Fair Use"

July 29, 2025

Two California district court judges recently issued competing rulings pertaining to fair use as a defense against the alleged improper use of copyrighted works to train large language models (LLMs). The two orders,[1] issued within days of each other, decided in the LLM owners' favor, finding the LLM owners met their burden of establishing "fair use" in connection with certain copyright claims brought by the plaintiffs. However, that is where the similarities end. The courts disagreed on the "most important" factor in fair-use analysis — the effect of the infringing use on the market for the copyrighted work.

The *Meta* court adopted a unique approach to this factor, looking at market dilution as opposed to direct market substitution. This approach was recently discussed in the Copyright Office's pre-publication of its "Copyright and Artificial Intelligence, Part 3: Generative AI" training report. The argument, however, was rejected by the *Anthropic* court, which also distinguished between copies used for training and non-training purposes, the latter category being further segregated between pirated and purchased data. (This note deals with copies for training purposes, unless noted otherwise.) However, because the defendant in *Meta* did not raise the "market dilution" argument, the court found, as in *Anthropic,* against the plaintiffs and for fair use.

The opinions, from the same federal district, will undoubtedly be reconciled and revised as the cases proceed and appeals play out, but they provide the first fulsome discussions concerning generative AI and fair use. They therefore provide initial indications of how courts will handle this novel and precedent-stretching technology and a potential roadmap for future cases and decisions.

Background

Both cases involve an LLM that was trained on copious amounts of data, including the plaintiffs' copyrighted works. Plaintiffs' central claim is that the LLM owners' copying of their works *for the use of training the LLMs* violated their copyright. Neither order substantively deals with alleged copyright violations for outputs — *i.e.*, chat responses — of the LLMs, as both courts acknowledge that neither set of plaintiffs claim substantial similarity of outputs to plaintiffs' works.

Defendants in both cases moved for summary judgment, arguing, *inter alia*, there were no issues of material fact concerning their respective fair use defenses.

Fair use is an affirmative defense to an otherwise valid copyright claim. The factors used to determine fair use are:

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.

4. The effect of the use upon the potential market for or value of the copyrighted work.

Both orders largely focus on the first and fourth factors.

Transformative Use vs. Market Effect

The judges both found in the LLM owners' favor regarding the first factor — holding that their use of copyrighted works to train LLMs was "transformative." The first factor looks at whether the new use "has a further purpose or different character," and, if so, whether this use is "transformative" and more likely to be "fair." *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 510 (2023). The *Anthropic* court found this use "transformative" in no uncertain terms: "the purpose and character of using copyrighted works to train LLMs to generate new text was quintessentially transformative." *Anthropic* at 13. The *Meta* court similarly found there was "no serious question that Meta's use" had a "further purpose" and "different character," in other words that the use was "highly transformative." *Meta* at 16.

The courts, however, disagreed on how to weigh this "transformative purpose" in their fair use analysis. The *Meta* court fastidiously argues the first factor is trumped by the fourth — market effect. As the court says: "Harm to the market for the copyrighted work is more important than the purpose for which the copies are made." *Meta* at 3 citing *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539,566 (1985). While the *Meta* court is correct that the fourth factor is frequently referred to as the "most important element of fair use," the court's analysis is admittedly more expansive than previous decisions.

The necessary premise of the *Meta* court's fourth-factor argument is that "indirect substitution is still substitution." *Meta* at 31. By widening this definition, the *Meta* court can presumably comply with the Supreme Court's dictate that the "only harm" that matters under this factor is "the harm of market substitution."[2] *Campbell v. Acuff-Rose Music, Inc*., 510 U.S. 569, 593 (1994). The alleged harm, according to the *Meta* court, is the LLM's ability to "enable the rapid generation of countless works that compete with the originals, even if those works aren't themselves competing." *Meta* at 28. In other words, the LLM is making it easier to produce "competing" works, and "by flooding stores and online marketplaces so that some of [the copyrighted] books don't get noticed and purchased, those outputs would reduce the incentive for authors to create." *Id.* at 31.

But the "indirect substitution" or "market dilution" argument, as the *Meta* court alternatively calls it, does not have much precedential backing. Indeed, the *Meta* court appeared to raise this argument on its own, as the defendants "barely g[a]ve this issue lip service." *Meta* at 4. In such a scenario, one would expect a court to remain silent or only raise the issue if it were to rely on black-letter precedent. Instead, the court rejected Meta's argument that such harm and the "market dilution" theory has "never made a difference in a case before," *not* by providing some precedent concerning market dilution, but by relying on a policy argument that fair use should take "account of 'significant changes in technology.'" *Meta* at 32.

The *Meta* court appears guided by (but did not cite to) the U.S. Copyright Office's recent guidance on the topic. The Copyright Office report endorsed this "market dilution" analysis concerning the fourth factor. While acknowledging the "uncharted territory" that generative AI occupies, the office — as echoed by the *Meta* court — found that "speed and scale at which AI systems generate content pose a serious risk of diluting markets for works of the same kind as in their training data." *Copyright and*

Artificial Intelligence Part 3: Generative AI Training, United States Copyright Office at 65 (May 2025). [3]

This rationale is indicative of the *Meta* court's focus on the radical and unprecedented technology before it. Generative AI and LLMs should be treated as a change in kind, not degree, and the *Meta* court rightly recognizes this change.

The *Anthropic* court, however, reached the opposite conclusion concerning the fourth factor (at least for the training copies).[4] There, the court analogized training an LLM to training a schoolchild to read and write on copyrighted books, which is "not the kind of competition or creative displacement that concerns the Copyright Act." *Anthropic* at 28. The *Meta* court takes direct aim at this "inapt analogy," arguing that "using books to teach children to write is not remotely like using books to create a product that a single individual could employ to generate countless competing works with a miniscule fraction of the time and creativity it would otherwise take." *Meta* at 3.

But *how* is it dissimilar? Is it the storage of the copyrighted material that is the difference? While an LLM owner must, almost out of necessity, store what an LLM is trained on, human memory is fickle and doesn't lend itself to such comparisons. If not storage, is it the sheer processing power of LLMs, *i.e.*, the "miniscule fraction" of the time it takes to teach an AI rather than a human that is the difference? The latter is certainly a compelling rhetorical point, although perhaps not as legally compelling. The former seems stronger legally but raises questions about human and artificial memory that would be difficult for any court to untangle. Thus, while at first blush it appears self-evident that "using books to teach children to write is not remotely like using books to create [an LLM]," closer scrutiny should be applied to this comparison — a scrutiny the *Meta* court did not provide.

Conclusion

These two rulings leave the fair-use landscape split between two incompatible rules of law. *Meta* emphasizes the market effects of the improper use, adopting a novel (and certainly persuasive) approach that rests on market dilution. This contrasts with *Anthropic*, which views LLM training as nothing more than giving a book to a schoolchild and teaching them to write — an approach that treats market dilution as irrelevant. These rulings will ripple through other industries beyond books, as parallel suits concerning music, images, and code are queued up in this space. Until further guidance is given — either by appeals courts or Congress — the takeaway from these rulings is a cautionary one for all involved: courts are willing to treat AI training as "fair use," but only on a fact-rich record. Litigants should be prepared to muster these facts and address the myriad of policy arguments that will be made.

We will continue to monitor this space, and please reach out if you or your company have any questions about these issues.

---

[1] The two cases are *Kadrey et al., v. Meta Platforms, Inc*., 23-cv-03417 (ND Cal.) and *Bartz et al., v. Anthropic* PBC, 24-cv-05417 (ND Cal.); the two summary judgment orders are ECF nos. 598 and 231, respectively. We will refer to these cases and orders as "*Meta*" and "*Anthropic,*" as appropriate.

[2] This is consistent with the Supreme Court's recent ruling, where the concurrence distinguishes between "whether consumers treat a challenged use 'as a market replacement' for a copyrighted work or a market complement that does not impair demand for the original." *Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 555 (2023) (Gorsuch, J. concurring).

[3] The Office continues, in an argument that would be perfectly at place in the *Meta* decision: "That means more competition for sales of an author's works and more difficulty for audiences in finding them. If thousands of AI-generated romance novels are put on the market, fewer of the human-authored romance novels that the AI was trained on are likely to be sold." *Id.* at 65.

[4] In fact — unlike the *Meta* court — the court in *Anthropic* was presented with this argument. *See Anthropic* at 28 ("Authors contend generically that training LLMs will result in an explosion of" competing works).

AUTHORS

Jennifer A. Kenedy

Glenn G. Pudelka

Jorden Rutledge

RELATED PRACTICES AND INDUSTRIES

Artificial Intelligence

Business Litigation

Intellectual Property