Kadrey v. Meta Platforms: A.I. Training and Market Harm to Authors

By Thomas A. Guerriero, PhD

Case Commentary | Independent Analysis | Published Thursday 19 November 2026

Independent commentary by a non-attorney for educational purposes; not legal advice and not filed with any court.

Case: *Kadrey v. Meta Platforms, Inc.*, No. 3:23-cv-03417-VC, Order Denying the Plaintiffs' Motion for Partial Summary Judgment and Granting Meta's Cross-Motion for Partial Summary Judgment, slip op. (N.D. Cal. June 25, 2025) (Chhabria, J.).

Background

Two days after *Bartz v. Anthropic* was decided in the same court, Judge Vince Chhabria ruled on whether Meta's use of books to train its Llama large language models was fair use. Thirteen authors, mostly well-known fiction writers, sued Meta for downloading their books from online "shadow libraries" and using them to train Llama. Both sides moved for partial summary judgment: the authors arguing that Meta's conduct could not be fair use, and Meta arguing that it must be fair use as a matter of law. Slip op. at 4.

The order records that Meta pursued licensing books for training, found it more difficult than anticipated, and turned to shadow libraries, first using one in October 2022, when it downloaded the Library Genesis database to test whether training on its contents was valuable. *Id.* at 10–11. The authors' claim that Meta also distributed their works through the torrenting process was not part of the summary judgment motions. *Id.* at 15, 40.

Meta won. But the order's reasoning is very different from its result, and that difference is the reason it matters.

Holding

The court granted summary judgment to Meta on its fair use defence to the claim that copying these plaintiffs' books for use as training data was infringement, and denied the authors' motion. *Id.* at 40. It stressed the limits of that ruling. The case was not a class action, so it affected only the thirteen authors. And the court said its 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." *Id.* at 5.

Reasoning

Market harm as the central question

The order opens with a general discussion before turning to the record. Generative AI, it says, has the potential to flood the market with endless amounts of competing works, and training on copyrighted works can create a product that will often dramatically undermine the market for those works. *Id.* at 1–2. It

acknowledges that training is highly creative and transformative, but states that "harm to the market for the copyrighted work is more important than the purpose for which the copies are made." *Id.* at 3.

The court directly criticised the reasoning in *Bartz*. It quoted Judge Alsup's comparison of training to "training schoolchildren to write well" and called it an "inapt analogy," reasoning that teaching children is not remotely like creating a product a single person could use to generate countless competing works with a tiny fraction of the time and creativity otherwise required. *Id.* at 3. It also rejected the argument that adverse rulings would halt the technology, observing that companies expected to earn billions from these products would find a way to compensate rights holders if training on copyrighted works were as necessary as they claimed. *Id.* at 3–4.

The four factors on this record

On the first factor, the court held that Meta's use was highly transformative because the purpose of the copying was to train language models that can generate diverse text and perform a wide range of functions. *Id.* at 16. Commercial use did not tilt the factor towards the authors. *Id.* at 19. Nor did Meta's use of shadow libraries give the authors an automatic win; the court said that calling the downloading "piracy" begs the question, since the point of the analysis is to decide whether the copying was unlawful. It treated possible bad faith as relevant at most, and found it did not move the needle on this record. *Id.* at 19–20.

The second factor favoured the authors because their books are creative works, but the court noted that this factor rarely plays a significant role. *Id.* at 23–24. The third factor favoured Meta: although it copied entire books, the amount was reasonable given the transformative purpose, since language models work better when trained on more high-quality material. *Id.* at 25.

Three theories of market harm

The fourth factor decided the case. The court described it as the single most important element and rejected Meta's suggestion that a strong first factor should effectively end the inquiry. *Id.* at 26. It then considered three ways training might harm the market for the authors' books.

The first theory was regurgitation. The order records that Meta post-trained its models to limit memorisation, and that Meta's expert, using a method designed to make models regurgitate training data, could get no model to generate more than 50 words and punctuation marks from the authors' books; the authors' own expert could get the best-performing Llama model to generate 50 such tokens in 60% of tests. *Id.* at 12. Llama could not reproduce enough text to matter. *Id.* at 4.

The second theory was a lost market for licensing books as training data. The court held that harm from losing fees for licensing a work for a transformative purpose is not cognizable, because otherwise the fourth factor would become circular and favour the rights holder in every case. *Id.* at 28.

The third theory was market dilution: that a model trained on books can generate large numbers of similar works that compete with the originals. The court called this the "potentially winning argument." *Id.* at 4. It reasoned that "indirect substitution is still substitution," so that a novel generated by a language model and bought instead of a human-written one would be substituting for it. *Id.* at 31. And because language models can generate millions of secondary works, no earlier fair use case had involved anything near the same potential to flood the market. *Id.* at 32.

The authors, however, barely raised dilution and presented no meaningful evidence on it. *Id.* at 4–5, 40. The court said its conclusion on factor four "may be in significant tension with reality" but was dictated by the authors' choice to put forward two flawed theories while failing to present meaningful evidence on the

market effects of training. *Id.* at 36. Because Meta's use was highly transformative, the authors needed to win decisively on the fourth factor; without evidence on dilution, that factor could only favour Meta. *Id.* at 40.

Significance

Kadrey and *Bartz* reached similar results on training but by very different routes. *Bartz* placed heavy weight on transformation and dismissed concerns about competing works. *Kadrey* treated market harm as the most important consideration and identified market dilution as a serious theory that the authors failed to support. The court's general prediction was explicit: in cases involving uses like Meta's, "it seems like the plaintiffs will often win, at least where those cases have better-developed records on the market effects of the defendant's use." *Id.* at 39.

The order also suggested how outcomes might differ. Copying to train a model for nonprofit purposes such as national security or medical research might be fair use even with some dilution, and plaintiffs whose works are unlikely to face meaningful competition from AI-generated ones may be unable to defeat fair use. *Id.* at 39.

The litigation continued after the ruling. Two days later, the court granted Meta summary judgment on the authors' claim under the Digital Millennium Copyright Act, reasoning that because the copying was fair use, removal of copyright management information could not have furthered an infringement. DMCA Order at 1 (N.D. Cal. June 27, 2025). On 25 March 2026, the court granted the authors leave to file a fourth amended complaint adding a contributory infringement claim based on Meta allegedly uploading works to torrenting networks while downloading them, and denied a motion for class discovery. Amendment Order at 1 (N.D. Cal. Mar. 25, 2026). A discovery order of 19 October 2026 addresses disputes in *Kadrey* together with related actions against Meta brought by Entrepreneur Media, Chicken Soup for the Soul, Cambronne and Cognella. Discovery Order at 1–2 (N.D. Cal. Oct. 19, 2026). This commentary asserts no later outcome on the distribution or contributory infringement claims.

Practical Takeaways

The following are general governance observations drawn from the decision, not legal advice about any particular use of data.

- **Market evidence decides cases.** The ruling turned on what the record showed about market effects, not on general views about AI.
- **Transformation is not a safe harbour.** The court treated a highly transformative use as defensible only absent evidence of significant market harm.
- **Dilution is now a live theory.** Organisations creating or licensing content should expect arguments about AI-generated substitutes to be developed with evidence.
- **Purpose and setting matter.** The court suggested nonprofit research uses and works unlikely to face AI competition might be treated differently.
- **Acquisition remains a separate issue.** Distribution through torrenting was left for separate proceedings.

These observations draw on pages 1–5, 11–12, 16–40 of the order and the later orders cited above. They are offered to support careful data governance and informed discussion with qualified counsel.

Open Questions

The order leaves central questions unresolved. It did not decide whether training is fair use where plaintiffs present strong evidence of market dilution. It did not decide whether Meta distributed the authors' books or whether such distribution could be fair use. Slip op. at 15, 40. And as a district court ruling limited to thirteen authors, it does not bind other courts.

For authors and for organisations that hold rights in written works, the most practical lesson of *Kadrey* is that the strength of a challenge to AI training may depend on evidence of how AI outputs actually affect demand for the original works.

References

- United States District Court for the Northern District of California, *Kadrey v. Meta Platforms, Inc.*, No. 3:23-cv-03417-VC, Order on Cross-Motions for Partial Summary Judgment (filed 25 June 2025), official U.S. Courts opinion published on GovInfo: https://www.govinfo.gov/content/pkg/USCOURTS-cand-3_23-cv-03417/pdf/USCOURTS-cand-3_23-cv-03417-37.pdf
- Same court, Order Granting Meta's Motion for Partial Summary Judgment as to the Plaintiffs' DMCA Claim (27 June 2025): https://www.govinfo.gov/content/pkg/USCOURTS-cand-3_23-cv-03417/pdf/USCOURTS-cand-3_23-cv-03417-38.pdf
- Same court, Order Granting Motion for Leave to File Fourth Amended Complaint; Denying Motion for Class Discovery (25 March 2026): https://www.govinfo.gov/content/pkg/USCOURTS-cand-3_23-cv-03417/pdf/USCOURTS-cand-3_23-cv-03417-41.pdf
- Same court, Discovery Order (19 October 2026): https://www.govinfo.gov/content/pkg/USCOURTS-cand-3_23-cv-03417/pdf/USCOURTS-cand-3_23-cv-03417-54.pdf
- Same court, *Bartz v. Anthropic PBC*, No. 3:24-cv-05417-WHA, Order on Fair Use (23 June 2025): https://www.govinfo.gov/content/pkg/USCOURTS-cand-3_24-cv-05417/pdf/USCOURTS-cand-3_24-cv-05417-0.pdf