



July 8, 2025

If you have any questions regarding the matters discussed in this memorandum, please contact the attorneys listed on the last page or call your regular Skadden contact.

---

This memorandum is provided by Skadden, Arps, Slate, Meagher & Flom LLP and its affiliates for educational and informational purposes only and is not intended and should not be construed as legal advice. This memorandum is considered advertising under applicable state laws.

---

One Manhattan West  
New York, NY 10001  
212.735.3000

525 University Ave.  
Palo Alto, CA 94301  
650.470.4500

## Fair Use and AI Training: Two Recent Decisions Highlight the Complexity of This Issue

Two recent summary judgment decisions out of the Northern District of California, issued only two days apart, highlight the complexity of deciding whether the unauthorized use of copyrighted works to train large language models (LLMs) is infringement or fair use.<sup>1</sup> In *Bartz v. Anthropic PBC* (Judge Alsup, June 23, 2025) and *Kadrey v. Meta Platforms, Inc.* (Judge Chhabria, June 25, 2025), authors alleged that their works had been used without permission to train LLMs<sup>2</sup>. In each case, the court found that, on the facts before it, the use of copyrighted works to train an AI model was highly transformative and fair use. Much of the press coverage has treated these holdings as sweeping endorsements of fair use, but each ruling is narrow and fact-specific.

We first provide the key points of these decisions and then a detailed summary of the opinions.

### Key Points

- **Decisions will be fact-specific.** Fair use decisions are always highly fact-specific. While the 40 or so “training data” cases filed to date are often lumped together, the *Bartz* and *Kadrey* decisions highlight how each case will need to be assessed on its specific facts, making it difficult to draw broad conclusions from any one case. While Judge Chhabria does not appear as sympathetic to the fair use argument as other jurists might be, even he notes, “tweak some facts and defendants [in another case] might win.”
- **Evidence of infringing outputs and market impact will be critical.** Many copyright owners were concerned that the *Bartz* and *Kadrey* complaints presented weak fact patterns to challenge fair use since plaintiffs presented no evidence of outputs generated by the AI models that infringed their works, and weak or no evidence of any market impact on their works. Each decision stresses the absence of any such evidence, and Judge Chhabria in *Kadrey* broadly states that “in most cases,” training LLMs on copyrighted works without permission is likely infringing and not fair use: “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.”

Indeed, Judge Chhabria expressed frustration at needing to reach a finding of fair use: “Given the state of the record, the Court has no choice but to grant summary judgment ... [A]s should now be clear, 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.”

---

<sup>1</sup> LLMs are a type of generative AI model that can generate text in response to user prompts.

<sup>2</sup> For background on an earlier motion to dismiss in *Kadrey*, please see our [Nov. 22, 2023, client alert](#).

# Fair Use and AI Training: Two Recent Decisions Highlight the Complexity of This Issue

## - LLMs present unique challenges from a fair use perspective.

Judge Chhabria makes the important point that LLMs are a unique technology and therefore present an unprecedented challenge from a fair use perspective. On one hand, in his view, training LLMs on copyrighted works is highly transformative (strongly favoring fair use), while on the other hand they could significantly dilute the market for a plaintiff's works (strongly favoring a finding of infringement). As Judge Chhabria notes, courts have never confronted a technology that is both so transformative yet so potentially dilutive of the market for the underlying works.

- **The role of indirect substitution.** In analyzing the fair use doctrine's "market impact" factor, Judge Chhabria in *Kadrey* endorses the concept of "indirect" market substitution; namely, that courts should take into account LLMs' unique ability to rapidly "flood the market" with content that serves as a substitute for the plaintiffs' copyrighted works. While Judge Chhabria acknowledges that indirect market substitution has typically not been considered by courts, he states that it is relevant for LLMs given their ability to create "literally millions of secondary works" in a "miniscule [*sic*] fraction of the time" it took to create the original works.

- **The market for licensing works for training.** Both Judge Alsup and Judge Chhabria rejected the argument that when considering the "market impact" factor of the fair use doctrine, courts should take into account the market impact on licensing that content to LLM developers. In their view, this argument is circular because a plaintiff is, in effect, arguing that it lost the opportunity to market its work for an activity that may be determined to be fair use and does not require a license.

- **Is "transformation" the end of the analysis?** Judge Chhabria made clear that even a strong finding of transformation, as was the case here, is not the end of the analysis: "There is certainly no rule that when your use of a protected work is 'transformative,' this automatically inoculates you from a claim of copyright infringement." In contrast, Judge Alsup seemed to be more persuaded by the transformation argument, which may have influenced the short consideration he gave to the market impact argument, as discussed below.

- **No thwarting of innovation.** Judge Chhabria rejected the argument that a ruling against LLM developers would thwart this nascent technology. Calling the argument "ridiculous," he notes that "these products are expected to generate billions, even trillions, of dollars for [developers]. If using copyrighted works to train the models is as necessary as the companies say, they will figure out a way to compensate copyright holders for it."

- **Going forward.** *Bartz* and *Kadrey* signal the beginning of the second phase of judicial decisions in the rapidly evolving field of AI-related copyright litigation. The initial phase consisted of motions to dismiss where the court was focused primarily on whether the plaintiffs had properly stated a claim. This current phase will consist of summary judgment motions where courts have an evidentiary record before them. The third phase will be various appellate reviews, and a likely fourth and final phase will be review by the Supreme Court.

## Background to *Bartz* and *Kadrey*

Both *Bartz* and *Kadrey* are cases brought by authors against LLM developers (Anthropic and Meta, respectively) alleging that these companies trained their models, in part, on the authors' content without permission, and thereby engaged in copyright infringement. In both cases, the plaintiffs alleged that their books were included in "pirate" or "shadow libraries" used by these companies for training. These libraries contain large volumes of content, generally without the authorization of the copyright holders.<sup>3</sup>

While the general allegations in the two cases are the same, there are some important differences between them. While *Bartz* is a putative class action brought on behalf of similarly situated authors, *Kadrey* is not and therefore concerns only the works of the plaintiff-authors. More importantly, the *Bartz* plaintiffs alleged not only that Anthropic engaged in copyright infringement in connection with training its AI model, but also that Anthropic infringed their works through the creation of a "central library of 'all the books in the world'" that would be retained forever. As discussed below, this additional claim was critical to Judge Alsup's decision.

## Comparison of the Fair Use Analysis

The manner in which Judges Alsup and Chhabria applied the four-factor fair use test highlights how different courts can analyze this complex issue and how different fact patterns can yield markedly different results.

<sup>3</sup> The *Kadrey* plaintiffs also alleged claims of direct and vicarious copyright infringement, removal of copyright management information (CMI) in violation of the Digital Millennium Copyright Act (DMCA), unfair competition and negligence under California law, and unjust enrichment. Other than the direct copyright infringement claim, the other claims were dismissed with leave to amend. The plaintiffs then amended their complaint to add a new direct infringement claim as well as claims under the DMCA and the California Comprehensive Computer Data Access and Fraud Act (CDAFA). The CDAFA claim was subsequently dismissed, and Judge Chhabria recently granted summary judgment against the plaintiffs on their DMCA claim.

# Fair Use and AI Training: Two Recent Decisions Highlight the Complexity of This Issue

---

## Purpose and Character of the Use

### Training AI Models

The “purpose and character” factor has, in recent years, often turned on whether the court deems the defendant’s use to be “transformative.” In both *Bartz* and *Kadrey*, the court unequivocally found that use of the plaintiffs’ copyrighted works to train the defendant’s AI models was transformative. Judge Alsup in *Bartz* described Anthropic’s use as “quintessentially transformative,” and Judge Chhabria in *Kadrey* highlighted that the purpose of Meta’s copying is to train its LLMs to perform certain functions while the purpose of the plaintiffs’ books is to be read by humans.

Judge Alsup and Judge Chhabria differed on how each step of the training process should be viewed. While Judge Alsup in *Bartz* found that each step in the process (*e.g.*, downloading, copying, etc.) needed to be considered on its own merits, Judge Chhabria viewed the downloading and training as one integrated process and that the initial downloading must be considered “in light of its ultimate, highly transformative purpose.”

### Anthropic’s Central Library

*Bartz* involved one additional use of the plaintiffs’ materials not found in *Kadrey*: Anthropic’s inclusion of these books as part of a “permanent, general-purpose library” that had potential uses beyond LLM training. This library included both purchased copies of the plaintiffs’ books as well as those obtained through “pirate” libraries. With respect to the purchased books, the court found that converting a print book to a digital file to save space and enable searchability was transformative, especially since the original print copy was destroyed. In contrast, the court held that including pirated copies of books in a permanent library for possible future use weighed against fair use.

### The Nature of the Copyrighted Work

With respect to the nature of the copyrighted works, both judges concluded that this factor favored the plaintiffs, since the books at issue are highly expressive works or contain expressive elements. Judge Chhabria rejected Meta’s argument that it used the plaintiffs’ books only to access their “functional,” not expressive, elements, stating that word order, word choice, grammar and syntax are expressive elements from which AI models learn about statistical relationships between words.

## The Amount and Substantiality of the Portion Used

### Training LLMs

Both judges found that Anthropic’s and Meta’s copying, respectively, of the plaintiff’s complete works was reasonably necessary in relation to the transformative purpose of training LLMs, especially given that the AI models would not output any meaningful amount of the plaintiffs’ works.

### Anthropic’s Central Library

Once again, Judge Alsup analyzed the use of the plaintiffs’ works to create its central library separately. He found that while this factor favors the defendant for the legitimately purchased copies, it favors the plaintiffs for the pirated copies.

### The Effect of the Use Upon the Potential Market for the Copyrighted Work

Judge Alsup and Judge Chhabria took divergent approaches when assessing the types of market harm that could arise from using copyrighted works to train AI models without permission.

### Market To License Works for AI Training

Both courts rejected the plaintiffs’ argument that the unauthorized use of their works harmed or precluded the market to license their works for AI training. As Judge Chhabria noted, this is a circular argument as one cannot claim they lost the ability to license their works for a use that was ultimately deemed a fair use.

### Direct Market Harm

In both cases, the court rejected the plaintiffs’ argument that the LLMs would directly harm the market for the plaintiffs’ books since there was no evidence the LLMs could regurgitate the plaintiffs’ books. However, in *Bartz*, Judge Alsup found that the pirated copies of the plaintiffs’ books that were used to build the central library could have been purchased and therefore created direct market harm.

### Indirect Substitution and Market Dilution

Judges Alsup and Chhabria took markedly different approaches as to whether the ability of LLMs to rapidly generate countless works that compete with the plaintiffs’ — even if those generated works are not themselves infringing — could satisfy the fourth factor through an argument of “indirect substitution.” Judge Alsup dismissed this argument out of hand, finding the argument analogous to the noncognizable harm of “training schoolchildren to write well,” which might also lead to more competition.

# Fair Use and AI Training: Two Recent Decisions Highlight the Complexity of This Issue

Judge Chhabria was more receptive to this argument, critiquing Judge Alsup's analogy as "inapt." Judge Chhabria noted that while historically "indirect substitution" was not recognized as a market harm, LLMs are unique in their ability to create "literally millions of secondary works" in a "miniscule [*sic*] fraction of the time" it took to create the original works. Judge Chhabria observed that "it seems likely that market dilution will often cause plaintiffs to decisively win the fourth factor — and thus win the fair use question overall." Nonetheless, Judge Chhabria ruled in favor of Meta on this factor because the plaintiffs failed to present any evidence of this type of harm, relying instead on mere speculation.<sup>4</sup>

Ultimately, both courts ruled that, on the facts before them, the use of the plaintiffs' books to train an LLM was fair use. In *Bartz*, Judge Alsup also found that downloading pirated books to create a permanent, general-purpose library was not a fair use.

---

<sup>4</sup> In many instances, Judge Chhabria expressed frustration with the plaintiffs for not making the correct arguments or presenting the proper evidence. On this point, he noted: "As for the potentially winning argument — that Meta has copied their works to create a product that will likely flood the market with similar works, causing market dilution — the plaintiffs barely give this issue lip service, and they present no evidence about how the current or expected outputs from Meta's models would dilute the market for their own works."

## What the Decisions Mean Going Forward

A critical factor in each of *Bartz* and *Kadrey* was the absence of any evidence that the LLMs could generate outputs that replicate or are substantially similar to the plaintiffs' works. These two cases therefore stand in contrast to a number of other pending LLM training data cases whose complaints include evidence of such replication. Repeatedly, in each decision, Judges Alsup and Chhabria comment on how they would have likely ruled differently if presented with any such evidence.

We can also expect that many plaintiffs will focus on evidence of indirect substitution to bolster their arguments of market harm when challenging defendants' fair use defense.

Finally, as noted earlier, we strongly expect that these decisions and future ones on these issues will get appealed, with potential petitions to the Supreme Court for review. We therefore remain in the early stages of this evolving legal issue.

---

## Contacts

### Stuart D. Levi

Partner / New York  
212.735.2750  
stuart.levi@skadden.com

### Pramode Chiruvolu

Counsel / Palo Alto  
650.470.4569  
pramode.chiruvolu@skadden.com

### Mana Ghaemmaghani

Associate / New York  
212.735.2594  
mana.ghaemmaghani@skadden.com

### Shannon N. Morgan

Associate / New York  
212.735.3711  
shannon.morgan@skadden.com

### MacKinzie M. Neal

Associate / New York  
212.735.2856  
mackinzie.neal@skadden.com