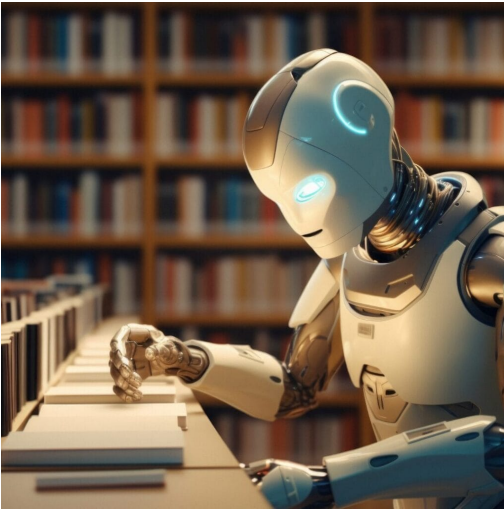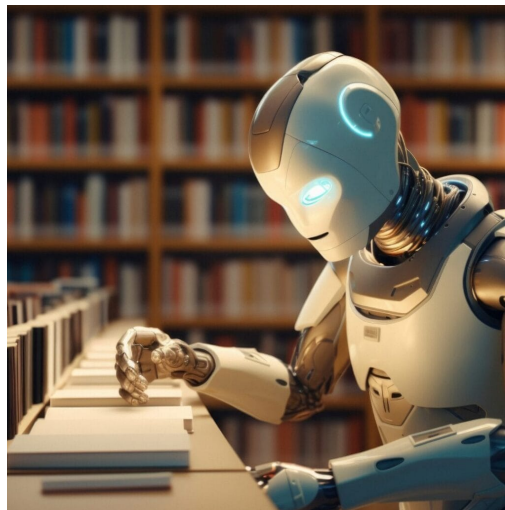# First Two Opinions Drop on AI and Copyright Fair Use

In the last week, two different federal district judges in the Northern District of California issued rulings in the first substantive decisions addressing whether use of extensive copyrighted material to train generative AI models is permissible fair use.

The court in the first case, [More](#)

In the last week, two different federal district judges in the Northern District of California issued rulings in the first substantive decisions addressing whether use of extensive copyrighted material to train generative AI models is permissible fair use.

The court in the first case, *Bartz v. Anthropic PBC*, lauded the defendant's large language model ("LLM"), " as "among the most transformative many of us will see in our lifetimes," and found that use of other parties books to train the models was a fair use (as long as the copies it used were not pirated).

The court in the second case, *Kadrey v. Meta Platforms, Inc.*, while also finding for the defendant on fair use, pointed to the plaintiff's failure to submit sufficient evidence of market harm from the copying – and opined that "in most cases" the defendant will lose, absent permission from or payment to the owner of the content used to train an AI LLM model.

Therefore, far from settling the issue, these two decisions may presage a battle between two approaches that courts may take to assess AI developers' fair use defense in the coming months and years.

Our detailed analyses and descriptions of each case follow.

Anthropic Decision

On June 23, Judge William Alsup in *Bartz v. Anthropic PBC* ruled on summary judgment that AI startup Anthropic did not engage in copyright infringement by training its LLM AI assistant, Claude, on published books.

Rather, Judge Alsup held that "the purpose and character of using copyrighted works to train LLMs to generate new text was quintessentially transformative" and constituted fair use. He likened Claude to an author in its own right: "Like any reader aspiring to be a writer, Anthropic's LLMs trained upon works not to race ahead and replicate or supplant them — but to turn a hard corner and create something different."

At the same time, while Judge Alsup held that digitization of physical books to create a digital library was a fair use in part because the format change itself was transformative, using *pirated* books to create the same library was not a fair use.

The court found that, among the books it used to train Claude, Anthropic knowingly copied and stored millions of pirated books (including from rogue websites) in a "central library" – and rejected the fair use defense with respect to use of these copies. Judge Alsup has scheduled a trial at the end of the year on the damages from copying the pirated books.

The *Anthropic* ruling suggests that AI companies may be developing the legal right to train LLMs on copyrighted works if they take care to avoid use of works known to be illegally copied to the internet.

It is equally, however, a reminder to companies training generative AI models to proactively source training material responsibly to try to avoid the potential consequences of a finding of willful copyright infringement, the resultant damages for which can be severe.

As the court noted: "That Anthropic later bought a copy of a book it earlier stole off the internet will not absolve it of liability," even if it might affect the extent of damages [JO1].

Kadrey Decision

On June 25, two days after the *Anthropic* decision, another Judge in the Northern District of California, Judge Vince Chhabria, came to a parallel conclusion on a similar set of facts – though reluctantly and for very different reasons – in *Kadrey v. Meta Platforms, Inc.*, finding in favor of Meta on cross motions for summary judgment on the issue of fair use.

In that case, Plaintiffs are thirteen (mostly famous) authors who sued Meta for downloading their books from online "shadow libraries" and using the books to train Meta's LLM generative AI models ("Llama").

Judge Chhabria, like Judge Alsup in *Anthropic*, found the use of generative AI "highly transformative" under the first statutory fair use factor in Section 107 of the US Copyright Act, the "nature and purpose" of the defendant's use.

However, Judge Chhabria was particularly focused on the fourth statutory factor, the effect of the use on the market for the plaintiff's work, noting: "harm to the market for the copyrighted work is more important than the purpose for which the copies are made."

Judge Chhabria characterized Judge Alsup's ruling as "brushing aside concerns" about the harm that generative AI can "inflict on the market for the work it gets trained on."

In his discussion of the fourth factor, Judge Chhabria highlighted that the relevant harm in that analysis was "market substitution." Though there could be other theories of market harm, he found the market dilution theory – that the Llama models could learn to generate works that were similar enough to compete with plaintiffs' own works and indirectly substitute for them – most compelling.

Judge Chhabria commented that "the plaintiffs' presentation [on market harm] is so weak that it does not move the needle, or even raise a dispute of fact sufficient to defeat summary judgment."

Crucially, he noted the absence of any argument by plaintiffs regarding market dilution – in the face of Meta's contention that plaintiffs had not presented any evidence that Meta's use of their books had harmed book sales. This factor therefore weighed in favor of Meta and tipped the scale in its favor.

Significantly, Judge Chhabria rejected plaintiffs' primary theory of market harm – that Meta's unauthorized use of their books to train the Llama models harmed the market for licensing their books for the same purpose. He held that, regardless of whether such a market exists, it is not one plaintiffs are "legally entitled to monopolize."

Judge Chhabria explained that, in every fair use case, the plaintiff suffers a loss of potential market share if the market is defined as the theoretical market to license the use at issue in that case. To prevent the fourth factor from becoming circular, "harm from the loss of fees paid to license a work for a transformative purpose is not cognizable."

However, he was equally hostile to Meta's argument that an adverse ruling would have a negative impact on the public interest because it would stop the development of LLMs and other generative AI technologies, asserting: "This is nonsense." Instead, he explained, "a ruling that certain copying isn't fair use doesn't necessarily mean the copier has to stop their copying—it means that they have to get permission" and, presumably, pay, for it.

In the introduction to his ruling, Judge Chhabria clearly limited his ruling to the parties in the case ("the ruling only affects the rights of these thirteen authors") and seemed to leave the door open for others to sue Meta for similar behavior in the future:"[T]his ruling does not stand for the proposition that Meta's use of copyrighted materials to train its language models is lawful."

"[I]n many circumstances it will be illegal to copy copyright-protected works to train generative AI models without permission," he wrote, "which means that the companies, to avoid liability for copyright infringement, will generally need to pay copyright holders for the right to use their materials."

There are a couple additional takeaways from reviewing he *Anthropic* and *Kadrey* decisions that were issued in rapid succession.

The judges seemed to agree that whether the copier used an authorized or pirated copy of the plaintiff's works is relevant to the fair use analysis, and that the role of "bad faith" in the fair use analysis remains in flux in the Ninth Circuit.

They disagreed, however, on whether courts should conduct an analysis of all stages of the copying at issue. Judge Alsup separately assessed the "initial, intermediate, and ultimate copies used by the copyist," while Judge Chhabria concluded that because Meta's "ultimate use of plaintiffs' books was transformative, so too was Meta's downloading of those books."

More detailed case summaries of Anthropic and Kadrey are set forth below.

<u>DETAILED DESCRIPTIONS OF THE CASES</u>

<u>*Bartz v. Anthropic PBC*</u> Case Summary

Defendant Anthropic is an AI software firm offering an LLM AI assistant called Claude that mimics human reading and writing. Plaintiffs Andrea Bartz, Charles Graeber, and Kirk Wallace Johnson (the "Authors") are human authors whose nine books Anthropic used, among millions of others, to train Claude. Along with two additional corporate plaintiffs, the Authors own all the copyrights in the nine books at issue.

Anthropic's goal was to use "all the books in the world" to train Claude. To reach that goal, Anthropic gathered millions of books, first by pirating books online, then by bulk purchasing print copies of books and scanning the books into digital form. Anthropic used only a subset of these books to train Claude, but kept the entire library as a permanent, general resource.

In short: "Anthropic used copies of Authors' copyrighted works to iteratively map statistical relationships between every text-fragment and every sequence of text-fragments so that a completed LLM could

receive new text inputs and return new text outputs as if it were a human reading prompts and writing responses."

Each work used for training was copied in four main ways. First, a working copy for the training set was created. Second, each work was "cleaned" to remove "lower-valued or repeating text (like headers, footers, or page numbers)." Third, each cleaned copy was translated into a "tokenized" copy according to an Anthropic-made dictionary. Fourth, each trained LLM retained a "compressed" copy of the work.

Notably, the Authors did not allege that Claude would provide to its users any of their works in full; they challenge the *inputs* to the LLMs, not the *outputs*.

Anthropic moved for summary judgment, arguing that its use of the pirated and physical books was fair use under Section 107 of the Copyright Act. At the outset, the parties disagreed about whether there were one or two "uses" at issue. Namely, whether Anthropic "used" the books only once, as it argues, to train Claude. Or twice, as the Authors argue, to first build a vast library of content and second to train LLMs using subsets of that content. Additionally, the Authors argued that uploading the print books to the digital format was an infringement.

Section 107 of the Copyright Act sets forth four non-exclusive factors to assess whether a given use of a copyrighted work is a permissible fair use: (1) the purpose and character of the use, including whether it is commercial, (2) the nature of the copyrighted work, (3) the amount and proportion of the work used, and (4) the effect of the use on the potential market for the original work.

Under the first factor, the court concluded that the copies used to train the LLMs were fair use. The Authors made three arguments against a finding of fair use: (1) using the works to train Claude was like using the works to train any person to read and write and the Authors should therefore be able to exclude Anthropic from this use, (2) the training was intended to memorize the works' creative elements, not just the non-protectable ones, and (3) computers should not be allowed to do what people do.

The court rejected all three arguments, reasoning that that the purpose and character of the use was "quintessentially transformative," likening the LLMs to a reader learning how to be a writer of new material.

The court's analysis differed as to the purpose and character of the use of the *purchased* books and the use of the *pirated* books to build the central library.

Turning first to the physical copies purchased in bulk, destroyed, and digitized, the court concluded that the change in format (which would ease storage and searchability) was transformative. The court cited a variety of examples considering changes in physical form to reason that storage and searchability were physical properties of the frame around the work or information properties, not creative properties.

While the court acknowledged Anthropic was a commercial entity and assumed it benefited from the format change, the court did not weigh that fact heavily because Anthropic did not create new copies of the works to share or sell. The court also considered that while Section 106(1) reserves to the copyright owner the right to make reproductions but concluded "no content was added or subtracted" in this format change.

As to the pirated books, the court firmly concluded that pirating copies to build a research library was not a transformative use. Indeed, it cast doubt on whether "any accused infringer could ever meet its burden of explaining why downloading source copies from pirate sites that it could have purchased or otherwise accessed lawfully was itself reasonably necessary to any subsequent fair use.

Anthropic made a number of arguments, including that because its ultimate use of the pirated books, training Claude, was highly transformative, all the stops along the way were also justified. But the court rejected that argument, citing the Supreme Court's decision in the *Warhol* case for the idea that courts must look past the "subjective intent of the user" to the objective use made of the work.

The court additionally noted that Anthropic's initial copies of the pirated books were not immediately transformed into their significantly altered form (training Claude) nor was every copy used for that purpose. As a result, the court concluded it should apply a line of decisions regarding intermediate copying that examine the "initial, intermediate, and ultimate copies used by the copyist" because the Ninth Circuit has held that intermediate copying may be infringing regardless of whether the end product also infringes those rights. Together, the court concluded all these points weighed against a finding of fair use.

The court easily found that the second statutory fair use factor, the nature of the copyrighted work, weighed against fair use in all instances because all the parties agreed the works contained expressive elements

On the third factor, amount and substantiality of the use, the court again split its analysis between the copies used to train the LLMs and the copies used to build the central library. Considering the copies used to train the LLMs, the court clarified that that the issue was not how much of the work was used in making the copy, but the amount and proportion of the work made accessible to the public. With this in mind, the court noted the lack of connection between Claude's outputs and the Authors' works and concluded that the copying was reasonable.

The Authors argued that that the copying was overly extensive and not strictly necessary, but the court rejected these arguments. Turning to the copies used to build the central library, the court found that for the print copies Anthropic purchased and digitized, it was entitled to keep the copies in a library so copying the entire work was what this purpose required and there was no "surplus copying." As a result, this factor favored fair use.

For the pirated copies, however, the court held that Anthropic lacked entitlement to the books in the first instance, so while its stated purpose was to train LLMs, its "objective conduct" was to seek out all the books in the world and retain them even after deciding it would not make further copies of them to train the LLMs. In that case, almost any unauthorized copying would have been too great and the third factor weighed against fair use.

Finally, the court considered the fourth statutory fair use factor, the effect of the uses on the market for the value of the Authors' works. As an initial matter, the court found that the copies used to train LLMs "did not and will not displace demand for copies of Authors' works, or not in the way that counts under the Copyright Act," because Claude did not reproduce their works – and therefore the factor weighed in favor of fair use.

The court rejected the argument that training LLMs would create an uptick in works competing with the Authors' works as well as the contention that a finding of fair use would displace an emerging market demand for licensing copyrighted works for the narrow purpose of training LLMs.

Against the latter point, Anthropic argued that the transaction costs of such a license would be so high it would cease developing its technology. While the court acknowledged it was required to assume Authors were correct on summary judgment because the record could support either account, it held that "such a market for that use is not one the Copyright Act entitles Authors to exploit."

In assessing the purchased copies of the works, the court concluded this factor was neutral because the court assumed that digitization displaced purchases of new digital copies that Anthropic would have made directly from the Authors, but yet again the court noted that "such losses did not relate to something the Copyright Act reserves for Authors to exploit" because it was a format change. The pirated copies, however, plainly displaced demand for Authors' books, copy for copy, and this factor therefore weighed against fair use with respect to the pirated books.

Putting all the factors together, the court concluded that the copies used to train LLMs were justified as a fair use because "[t]he technology at issue was among the most transformative many of us will see in our lifetimes." The copies used to convert purchased print copies into digital library copies were justified for a different fair use. The pirated copies used to build a library were not a fair use. The court did not grant summary judgment as to the copies made from the central library but not used for training. The court ordered a trial on the pirated copies used to create Anthropic's central library and the resulting damages.

Kadrey v. Meta Platforms, Inc. Case Summary

Plaintiffs are thirteen authors, mostly famous fiction authors, who hold copyrights in various books. They sued Defendant Meta for downloading their books from online "shadow libraries" and using the books to train Meta's LLM generative AI models ("Llama").

The plaintiffs originally brought claims for direct and vicarious copyright infringement, removal of copyright management information in violation of the DMCA, unfair competition, unjust enrichment, and negligence. All the claims except the direct copyright claim were dismissed early on. Plaintiffs then amended to expand their copyright claim to include a theory of infringement by distribution, and to add a different DMCA claim and a claim under the California Comprehensive Computer Data Access and Fraud Act which was also dismissed. The parties filed cross motions for summary judgment on the question of fair use.

Plaintiffs offered two primary theories for how the markets for their works are affected by the unauthorized use of their works to train an AI model. First, they contended Llama can reproduce small chunks of text from their books and Meta, by using these works to train Llama without permission, has diminished their ability to license these works for this very purpose. Second, they argued that Meta has copied their works to create a similar product that could create market dilution.

The court noted that fair use is a "flexible concept" and that the four statutory factors listed in Section 107 of the US Copyright Act are not exhaustive. Quoting *Warhol*, the court emphasized that the fourth factor – the effect of the defendant's use on the potential market for the original work – is "undoubtedly the single most important element of fair use" and explained that the secondary user will likely have difficulty meeting its burden without "favorable evidence about relevant markets."

While the rightsholder need not prove or provide evidence about relevant markets, they "may bear some initial burden of *identifying* relevant markets." The court also noted that where a fair defense use fails, the defendant need not always stop use, it may be able to negotiate and pay for a license to account for the fact that the conduct will otherwise harm the market for the original work.

The court provided an overview of generative AI and large language models and explained that books are particularly valuable training data for LLMs, because they allow the LLMs to work with large amounts of text at once to train the LLM's "memory." Meta previously used free website data, metadata, and text to train the Llama models. It also attempted to enter licensing agreements with publishers, but the rights to license books for AI training are often held by individual authors, not publishers.

Eventually, Meta turned to "shadow libraries," an online repository of books, academic journal articles, music, films, etc., available online for free regardless of whether the media is copyrighted. To download the materials more quickly, Meta torrented them. Torrenting, the court explained, is a filesharing technique that involves simultaneous upload and download of distribution of small portions of larger files from many different sources. The parties dispute to what extent Meta uploaded the data it torrented.

Once Meta had the books, it post-trained the Llama models to prevent them from "memorizing" and outputting text from training data, including copyrighted material.

The court considered each of the Section 107 factors in turn. First, the court found that that the character and purpose of the use factor favored Meta because it was "highly transformative." It reviewed various tasks the Llama models can perform and contrasted them against the purpose of plaintiffs' books, which is to be read.

The court rejected an argument in an amicus brief that Meta's use has the same purpose and character as the books because an LLM training on a book is akin to a human reading one in the same way that a professor copies a book and gives it to a student. The court reasoned that an LLM ingests text to learn patterns and copied the books to create tools.

The court also rejected plaintiffs' argument that the use was not transformative because the Llama models have "no critical bearing" on their books. The court pointed out that critique and commentary are not the only uses that qualify as fair use. Last, the court rejected the argument that the Llama models only repackage the plaintiffs' books, explaining that, as noted above, the Llama models are specifically trained not to generate long portions of text that would function as repackaging.

Under the first factor, the court also considered the commercial nature of Meta's use. Although the Llama models are available under a free license, they were developed for commercial reasons and Meta expects them to generate hundreds of millions to billions of dollars.

While the court disagreed with Plaintiffs that the fact that Meta pirated the books automatically means there cannot be fair use, it also disagreed with Meta that its use of the shadow libraries was irrelevant to the analysis. As in *Bartz*, the court noted that the law regarding bad faith seems to be in flux. The court additionally noted that downloading copyrighted material from shadow libraries would be relevant if it benefitted the libraries and perpetuated unauthorized copying.

The court's last consideration under the first factor was the relationship between Meta's downloading of plaintiffs' books and Meta's use of the books to train the Llama models, where it determined that the downloading should be considered in light of the ultimate transformative purpose.

The court held that the second factor favored plaintiffs because their books are highly expressive works, rejecting Meta's argument that it only used plaintiffs' books for their functional elements, not their creative expression. In doing so, the court explained that the "statistical relationships" that the Llama models aim to learn are the product of creative expression.

Considering the third factor, the court noted that it wasn't particularly relevant in this case because the Llama models won't output a meaningful amount of plaintiffs' works. Nonetheless, the court held that the factor favored Meta even though it copied the books in their entirety, because the amount was reasonable given its relationship to Meta's transformative purpose.

Finally, the court considered the fourth factor, looking to both the extent of market harm from Meta's activity and whether unrestricted and widespread similar conduct would "result in a substantially adverse impact on the potential market for the original."

The relevant harm in this analysis is market substitution. The court went on to propose three ways a plaintiff might try to argue market harm in the case of generative AI: (1) the model will regurgitate their works and allow users to access them for free, (2) the market for licensing their works to train generative AI is diminished, and (3) the model can generate similar works to compete with the originals and indirectly substitute for them.

The court considered the third to be the rationale most relevant to the effect on the market factor, and found the plaintiffs failed to proffer relevant evidence on this point. And while fair use is an affirmative defense, defendant introduced evidence of lack of market harm and plaintiff failed to counter the showing, so the fourth factor weigh in defendant's favor.

The court also briefly considered: (1) whether Meta's use of shadow libraries benefitted those libraries, but ultimately found that plaintiffs had made no such showing, and (2) the public benefit associated with Meta's copying, which it found to slightly favor Meta because the Llama models will help users generate new creative outputs.

 [JO1]I have seen some analysis suggesting that this case establishes that obtaining copyrighted materials through piracy cannot be excused by subsequent transformative use, but I think that's wrong because the LLMs were trained from subsets of the central library which contained pirated material. Am I understanding that right? I think the intermediate copying section supports my reading that the central library was an initial copy and the training was the ultimate copy.