IP Alert | Two Big AI/LLM Copyright Rulings: Fair Use OK, But Piracy/Output Still Matters

By Victoria Webb and Kirk Sigmon

This week, two significant copyright decisions relating to artificial intelligence ("AI"), machine learning training, and fair use were issued from the Northern District of California. In the first case, Bartz et al. v. Anthropic PBC, No. 24-cv-05417, Doc. No. 231 (N.D. Cal. June 23, 2025) ("Anthropic"), the Court issued an order on fair use indicating that the training of Large Language Models ("LLMs") using legally-acquired copyrighted materials constituted fair use, but that copyright owners might still have a right to sue for infringement if those materials are pirated. In the second case, Kadrey et al. v. Meta Platforms, Inc., No. 23-cv-03417, Doc. No. 598 (N.D. Cal. June 25, 2025) ("Meta"), the Court similarly found that the training of machine learning models using copyrighted materials could be fair use, but indicated that the inquiry of fair use is highly factually specific and could result in a different outcome based on (for instance) whether the resultant model could substantially replicate the input copyrighted materials. While the fair use factors remain set forth in Section 107 of the U.S. Copyright Act, the law applying these factors and the fair use defense in the AI context is in flux (we expect further legal developments, including appeals, to follow). So these initial orders help provide a rough idea of preliminary best practices for companies using data to train artificial intelligence models.

Anthropic **Order: Training Using Lawfully-Acquired Content Fair Use, But Does Not Excuse Piracy**

The Court's summary judgment order in Anthropic pertains to Anthropic's use of various copyrighted books as part of training Anthropic's LLMs, including its "Claude" AI software service. As noted by the Court, Anthropic's LLMs were trained using books from two different sources: one volume of books was downloaded (that is, pirated) via the Internet, whereas another volume of books originated from digital scans of physical books that Anthropic purchased in bulk. Regardless of origin, Anthropic performed various processing steps to prepare those books for LLM training, assembled the digital copies into a "permanent" and "central library," and ultimately used those copies to train various LLMs. Various authors of the books sued Anthropic for copyright infringement. Anthropic quickly moved for summary judgment on its fair use defense, asserting that its use of those authors' books was necessary for training LLMs.

Considering various factors of fair use, the Court in Anthropic ultimately found that the training of an LLM using copyrighted material was, standing alone, fair use under Section 107 of the U.S. Copyright Act. The Court characterized use of copyrighted works to train

LLMs (but not as otherwise stored for other purposes) as "exceedingly" and "quintessentially transformative," analogizing such behavior to the process via which human beings learn through reading others' works. The Court was also quite deferential to Anthropic's arguments that it needed to use a large quantity of books, noting the need for "monumental" amounts of data to train LLMs.

That said, the Court's summary judgment order contains a key caveat: that training of an LLM constitutes fair use here does not excuse the piracy of copyrighted materials collected to perform that training. Noting that there is "no carveout . . . for AI companies" in the Copyright Act, the Court criticized Anthropic's piracy as "inherently, irredeemably infringing even if the pirated copies are immediately used for the transformative use [of training LLMs] and immediately discarded." In other words, Anthropic did not have the right to create a "central library" full of pirated content for nebulous "research" purposes. The Court contrasted these pirated materials with the books that Anthropic legitimately purchased (and then broke apart and scanned), characterizing Anthropic's transformation of those works into a digital form as "transformative under fair use" (and seemingly praising the fact that Anthropic destroyed physical copies after they were scanned). The Court also suggested that there might be limited circumstances where such piracy might be permissible (e.g., where copies were wholly unavailable for purchase/legal acquisition and/or where piracy was otherwise "necessar[y]" in some sense) and/or circumstances where Anthropic could have acquired the books for free (e.g., by "borrowing" copies of books from a reference library), but those circumstances did not apply to Anthropic's pirated copies. The Court further noted that Anthropic could not cure its infringement by later buying "a copy of a book it earlier stole" via the Internet. In short, the Court refused to allow Anthropic to "steal a work [it] could otherwise buy" just for the purposes of the purported "fair use" of LLM training. So the case will now continue to trial on what the Court estimates to be over 7 million "pirated copies [of books] used to create Anthropic's central library and the resulting damages, actual or statutory (including for willfulness)." Trial is currently set for December 2025, and notably, damages could be quite steep with statutory damages starting at $750 per book.

Critically, the Court's summary judgment ruling in Anthropic applies only to training data, not output from a trained machine learning model. As noted by the Court, "no output to the public was even alleged to be infringing." In other words, the Anthropic Court was not presented evidence that Anthropic's LLMs could produce output of "any exact copies [or] even infringing knockoffs" of the authors' works, and indicated that the authors "remain free to bring that case in the future."

Meta **Order: Training is Fair Use, But Might Depend on Output/Other Considerations**

The Court's ruling in Meta pertains to similar factual circumstances as those in Anthropic. Various authors sued Meta for allegedly downloading their works unlawfully and using them to train machine learning models (such as Meta's "Llama" software platform). Indeed, the Court seemed to openly assume based on the facts presented that Meta downloaded unauthorized (e.g., pirated) copies of various copyrighted works, although there was some dispute regarding whether Meta facilitated others' piracy of those books via the BitTorrent file sharing protocol. Meta, like Anthropic, filed for summary judgment on the issue of fair use. With that said, unlike in Anthropic, the authors in Meta also argued that Meta's

models were capable of reproducing "small snippets of text from their books"—that is, the authors' arguments seemed to suggest that both the training of the model and the output of the model could be infringing.

Somewhat like the Court in *Anthropic*, the Court in *Meta* generally found that Meta's use of the authors' books for training Meta's LLM had a "further purpose" and "different character" than the books, and was "highly transformative." This was, in no small part, because LLM training is different from how a person reads a book: for instance, the process of using a book for LLM training was generally not analogous to a user simply reading the book for edification/enjoyment. Moreover, the Court placed great weight on the fact that Meta's Llama model was generally unable to generate "more than 50 words from any of the plaintiffs' books," noting that this "does not threaten" the market value of the authors' books.

That said, the Court in *Meta* deviated from the *Anthropic* decision, explaining that the *Anthropic* decision "focused heavily on the transformative nature of generative AI while brushing aside concerns about the harm it can inflict on the market for the works it gets trained on." And although the authors argued that Meta's activities could provide market dilution vis-à-vis the output of their LLMs competing with their works, the Court characterized these as "half-hearted argument[s]" given that Meta's models were unable to reproduce the works. The Court in the *Meta* decision also indicated that the record was insufficient to definitively rule that the fair use defense did not apply to Meta's piracy of the authors' books and/or that the output of Meta's models was infringing. In other words, the *Meta* opinion suggests that there may be circumstances where an AI company could freely download copyrighted works insofar as their use of those works was to train an LLM that would ultimately be incapable of creating "competing works."

Given this, the readers should be wary: the Court was careful to couch much of its analysis on the particular facts and arguments before it in *Meta*, noting that tweaks to the facts might render remarkably different conclusions. The Court in *Meta* seemed overall reluctant in reaching its fair use conclusion, noting it "ha[d] no choice but to grant summary judgment" here "[g]iven the state of the record," and explaining that "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."

Legal Trend thus Far: Training is Fair Use, Piracy Still Dangerous

Both the order in *Anthropic* and the order in *Meta* generally suggest that the training of machine learning models using legally-acquired copyrighted works is, standing alone, transformative and thus highly likely to be found fair use. In other words, insofar as other forms of copyright infringement are not found (e.g., if a company legally acquires books and uses them for training), it is highly likely that courts will find such training to be fair use.

With that said, the differences between *Anthropic* and *Meta* underscore the risks of (1) piracy to acquire books and (2) designing models (e.g., LLMs) that can substantially reproduce and/or compete with copyrighted works. In particular, the *Anthropic* case suggests that, even if training is fair use, copyright owners might pursue action against entities that pirate their works for the initial act of piracy itself. Moreover, the Court's discussion in *Meta* suggests that, even where copyrighted works are legally acquired and

used to train a model, there may be no fair use where that trained model is capable of substantially reproducing the input copyrighted works or otherwise impacting the market for copyrighted works.

Conclusion: The Current "Gold Standard" is Legally-Acquired Training Data for Transformative Purposes

Given Anthropic and Meta, those developing AI models of any sort (LLMs, machine learning models, or the like) should seek to (1) use legally-acquired training data to (2) generate models that cannot substantially reproduce their training data. In other words, as suggested by the Court in Meta, the "gold standard" might involve the use of copyrighted books "to train an LLM for nonprofit purposes" unrelated to those books, such as for "national security" or "medical research" purposes. All the same, we anticipate that the law in this area will remain in flux (and subject to various appeals) for some time, so additional caution is always warranted.

Posted: June 26, 2025