AI Model Training is Unfair – In Limited Circumstances

A Delaware federal district court made headlines this week by issuing the first court decision rejecting fair use as a defense in training artificial intelligence (AI) models with copyrighted content. In *Thomson Reuters Enterprise Centre GMBH v. Ross Intel. Inc.,* Judge Bibas held, among other things, that two of the four factors analyzed in determining whether copyright infringement is defendable as fair use weighed heavily in favor of the copyright owner – in this case, Thomson Reuters. *Thomson Reuters Enterprise Centre GMBH v. Ross Intel. Inc.,* No. 1:20-cv-613-SB, 2025 U.S. Dist. LEXIS 24296 (D. Del. Feb. 11, 2025).

Procedural Background:

Almost as show stopping as the Court's decision itself, the path in which the Court entertained the issue of fair use is captivating. It is not every day that a federal court judge *sua sponte* endeavors reconsideration of their prior decision on summary judgment, especially on the eve of trial. But that is exactly what Judge Bibas did in this case. Specifically, the Court had already issued its summary judgement decision in 2023, largely denying Thomson Reuter's motions for summary judgment on copyright infringement and the fair-use defense. But as trial drew near (scheduled for August 2024), Judge Bibas "studied the case materials more closely and realized that [his] prior summary-judgment ruling had not gone far enough." *Thomson Reuters v. Ross* at 3. Therefore, the Court adjourned trial and invited the parties to renew their summary judgment briefing.

Thomson Reuters sought summary judgment of copyright infringement and denial of any related defenses including fair use and Ross sought summary judgment of non-infringement or in favor of fair use. As a result of the parties' supplemental briefing, the Court revised its original decision on summary judgment. In particular, the Court awarded partial summary judgment to Thomson Reuters, not Ross, on direct copyright infringement and related defenses. While those portions of the Court's opinion are informative, for brevity this article will focus on the Court's denial of Ross's fair use defense. If you would like to review the court's full opinion, it is available here.

Ross's Use of Thomson Reuter's Headnotes:

By way of background, Thomson Reuters owns the well-known legal research tool – Westlaw. Defendant Ross Intelligence is a company that created a legal-research search engine that uses artificial intelligence to compete with Westlaw. *See id.* at 3. In order to train its AI search tool, Ross needed a database of legal questions and answers, which led Ross to seek a license to Westlaw's content, to which Thomson Reuters refused. Undeterred, Ross entered into a contract with LegalEase to obtain training data in the form of "Bulk Memos." Bulk Memos are essentially compilations of legal Q&A developed by LegalEase's employee lawyers. LegalEase provided its lawyers with a guide explaining how to create those questions using Westlaw headnotes. LegalEase sold Ross approximately 25,000 Bulk Memos which Ross used to train its AI search tool. When Thomson Reuters learned of this, it sued Ross for copyright infringement. Notably, this case relates to non-generative AI, and was filed in 2020, well before ChatGPT arrived on the AI scene.

This is significant in that the Court's decision may be informative with respect to the analysis of whether use of copyrighted material to train generative AI tools may constitute fair use, but it is by no means dispositive of this query. Something to keep in mind as we continue to track developments in *The New York Times v. OpenAI and Microsoft case*.

Why Isn't Ross's AI Training a Fair Use?

Courts consider four factors as it analyzes whether use of a copyrighted work is defendable as a fair-use and the alleged infringer bears the burden of proof on this defense. The following factors carry varied weight: (1) the use's purpose and character, including whether it is commercial or not for profit; (2) the copyrighted work's nature; (3) how much of the work was used and how substantial a part it was relative to the copyrighted work's whole; and (4) how the use affected the copyrighted work's value or potential market. See 17 U.S.C. Section 107(1)-(4). The first and fourth factors weigh most heavily in the analysis. *Thomson Reuters v. Ross* at 16, *quoting Authors Guild v. Google, Inc.,* 804 F.3d 202, 220 (2d Cir. 2015). While the Court noted that fair use is a mixed question of law and fact, in the instant case the Court held that once copying is determined, the remaining issues for consideration are not related to historical fact, intent or factual prediction, rather it is about how to apply the law to the facts, which is a question for the judge not the jury. *Id.*

First, Judge Bibas considered Ross's purpose and character of use, which necessarily requires an analysis of whether it was commercial, and whether the use was transformative. *Id. citing Andy Warhol Found. For the Visual Arts Inc. v. Goldsmith,* 598 U.S. 508, 529-31 (2023). Because Ross and Thomson Reuters used the Westlaw headnotes for "very similar purposes and Ross's use is commercial, this factor disfavor[ed] fair use." Id. While Ross's use is obviously commercial, the Court spent a bit more time explaining why it was not transformative. Whether use is "transformative" is essentially related to the purpose of the use. "If an original work and a secondary use share the same or highly similar purposes, and the second use is of a commercial nature, the first factor is likely to weigh against fair use absent some other justification for copying." *Warhol* at 532-33. The Court held that Ross's use was not transformative because it did not have a "further purpose or different character" than Thomson Reuters's. *Id.* at 529.

Here, the record established that Ross used Westlaw headnotes as AI data to create a legal research tool to compete with Westlaw. When a user enters a legal question, Ross's tool responds with relevant judicial opinions that have already been written. This is strikingly similar to the process in which Westlaw uses headnotes.

In an effort to justify its copying as a fair use, Ross argued that the headnotes do not appear as part of the final product that Ross provides its consumers and that any copying occurred at an intermediate step: Ross turned the headnotes into numerical data about the relationship among legal words to feed into its AI model. Ross's argument is rooted in case law addressing fair use in analyzing computer programs[1], but the Court distinguished this argument. Specifically, the Court held that copying computer code is not the same as copying written words because it is well settled that "computer programs differ from books, films, and other literary works in that such programs almost always serve functional purposes." *Google,* 593 U.S. at 21. Second, the computer programing cases rely on a factor that is not at play here, in that "copying was necessary for competitors to innovate," or at least that is with Judge Bibas held in this particular decision. *Thomson Reuters v. Ross* at 18. I state it this way because again, this is a first decision in what could be the beginning of years' worth of court precedent on AI copyright cases. One might argue that copying of data to train AI models is "necessary for competitors to innovate." Perhaps that is not the case in this particular dispute, but a door is open to many alleged infringers in

---

1 See e.g., *Google LLC v. Oracle Am., Inc., 593 U.S. 1, 30-32 (2021)*; *Sony Comput. Ent., Inc. v. Connectix Corp., 203 F.3d 596, 599, 606-07 (9th Cir. 2000)*; and *Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1514-1515, 1522-23 (9th Cir. 1992)*

district courts around the country struggling to defend their business model and AI tools.

While the Court addressed the other two factors the analysis was of no moment and had no impact on the Court's ultimate decision finding in favor of the copyright holder and holding that use of copyrighted material to train AI models does not constitute fair use – at least in this case.

Takeaways:

The Court's holding in this case is narrowly tailored in many ways, but it likely will result in an uptick in copyright infringement claims against AI tools by content creators. One thing is certain and that is that this Copyright meets AI legal landscape is just starting to take shape, with much unknown terrain yet to be explored.

For more information on this case or the AI copyright landscape, contact Jessica Copeland or any attorney in the artificial intelligence or intellectual property practices at Bond with whom you regularly conduct business.