



court unplugs ai fair use defense, but context is key

Mark Humphrey
MSK Client Alert

February 12, 2025

Yesterday, the United States District Court for the District of Delaware became the first court in the United States to issue a **substantive decision** on whether using copyrighted material to "train" an artificial intelligence (AI) tool is protected by the fair use doctrine, finding that fair use did not apply under a rather unique set of facts. *Thomson Reuters Enterprise Centre GMBH and West Publishing Corp. v. Ross Intelligence Inc.*, Case No. 1:20-cv-613-SB ("Order"). While not addressing generative AI, the opinion considers legal issues that future courts will inevitably confront when deciding cases involving generative AI tools and technologies.

Factual and Procedural Background

Thomson Reuters owns the Westlaw legal research platform. In addition to case law, statutes, and treatises, Westlaw offers paid access to legal research aids, including a system of annotations known as "headnotes" which summarize key points of law and case holdings. Headnotes are organized within Westlaw using a feature called the "Key Number System," which the court in this case described as "a numerical taxonomy" of legal issues and concepts. Order at 2. As the court noted, Thomson Reuters owns copyrights in Westlaw's copyrightable material.

Ross Intelligence Inc. ("Ross") sought to create its own legal research search engine using artificial intelligence. But to "train" its AI search tool, Ross needed a database of legal questions and answers, and therefore had discussions with Thomson Reuters about licensing Westlaw content. It is important to note that Ross's AI tool is a search engine that does not generate new content in response to user prompts. Thus, unlike the tools at issue in other pending AI litigation (e.g., ChatGPT, Claude, etc.), Ross's tool was not considered *generative* AI.

attorneys

Mark C. Humphrey

practice areas

Artificial Intelligence (AI)
intellectual property
litigation
specialties



court unplugs ai fair use defense, but context is key

Unwilling to aid a competitor, Thomson Reuters declined to provide a license. Undeterred, Ross worked with a company called LegalEase to get training data in the form of "Bulk Memos," which the court's opinion described as "lawyers' compilations of legal questions with good and bad answers." Order at 3. To create Bulk Memos, LegalEase gave lawyers a guide explaining how to create questions using Westlaw headnotes, but clarified that the lawyers should not just copy and paste the headnotes. Ross purchased approximately 25,000 Bulk Memos from LegalEase, which it used to train its AI search tool. "In other words, Ross built its competing product using Bulk Memos, which in turn were built from Westlaw headnotes." *Id.*

Thomson Reuters filed the lawsuit in 2020. In November 2023, the court denied Thomson Reuters's summary judgment motions, brought on copyright infringement and fair use, and set the case for a trial to begin in August 2024. Shortly before trial was to begin, the court continued the trial and invited renewed summary judgment motions. Thomson Reuters moved for partial summary judgment on direct infringement and related defenses, while Ross moved for summary judgment on Thomson Reuters's copyright claims. Both parties moved for summary judgment on fair use.

Copyrightability of Thomson Reuters's Westlaw Headnotes and Key Number System

A key issue on which the court's prior order denied summary judgment was whether the headnotes and Key Number System were sufficiently original to be protectable by copyright. The court had previously found that originality depended on how much the headnotes overlapped with the non-copyrightable text of the judicial opinions, and had held that the copyrightability of the Key Number System was a jury question because of Ross's allegation that "most of the organization decisions are made by a rote computer program and the high-level topics largely track common doctrinal topics taught as law school courses." In the new opinion and Order, the court cited the seminal decision in *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345 (1991), which recognized the "extremely low" threshold for originality that requires "some minimal degree of creativity...some creative spark." The court noted that the key question is whether a work is original, not how much effort went into developing it.

The court went on to find that the headnotes "easily clear[] [the] low bar" of originality because they constitute an original, factual compilation, as Thomson Reuters made "choices as to selection and arrangement" using "a minimal degree of creativity." Order at 7. The court also found that each headnote constituted an individual, copyrightable work. Analogizing to a sculptor chiseling a block of raw marble, the court found that "even a headnote taken verbatim from an opinion is a carefully chosen fraction of the whole[,] and said that "[i]dentifying which words matter and chiseling away the surrounding mass expresses the editor's idea about what the important point of law from the opinion is." *Id.* In the court's view, "[t]hat editorial expression has enough 'creative spark' to be original." *Id.* at 7-8. Ultimately, the court made a sweeping ruling that "all headnotes, even any that quote judicial opinions verbatim, have original value as individual works." *Id.* at 8.

The court also found that the Key Number System could be an original compilation regardless of whether it was "novel"; instead, it merely needed to be "independently created by" Thomson Reuters. Order at 8. The court further found that "[t]here are many possible, logical ways to organize legal topics by level of granularity[,] and that "[i]t is enough that Thomson Reuters chose a particular one." *Id.*



court unplugs ai fair use defense, but context is key

Copying of Original Elements – Actual Copying and Substantial Similarity

The court then turned to whether Ross had engaged in actual copying (which requires a finding of access and probative similarity), and whether the headnotes and Bulk Memos were substantially similar in protected expression. As the court framed it, the dispute boiled down to whether the Bulk Memo questions copied Thomson Reuters's headnotes, or were instead taken from uncopyrightable judicial opinions. The court thus considered a batch of 2,830 headnotes identified by Thomson Reuters and compared them against the Bulk Memo questions and judicial opinions.

The parties agreed that LegalEase had access to Westlaw and used it to create the Bulk Memos. With regard to probative similarity, while noting that Ross's expert witness appeared to concede in her expert report that the 2,830 headnotes were created by actually copying Westlaw headnotes, the court compared the headnotes against the Bulk Memos and the judicial opinions. The court then granted summary judgment for Thomson Reuters on actual copying with respect to 2,243 of the headnotes, specifying that with respect to these particular headnotes, "actual copying is so obvious that no reasonable jury could find otherwise."

On substantial similarity of protected expression, the court applied a concept underlying the Ninth Circuit's view on "thin" copyrights, as well as the Second Circuit's "more discerning" ordinary-observer test: "The less protectable expression a work contains, the more similar the allegedly infringing work must be to it." Order at 14. The court then granted summary judgment on substantial similarity for Thomson Reuters on the same 2,243 headnotes, because it found that for those headnotes, substantial similarity again "is so obvious that no reasonable jury could find otherwise." The court noted that it was granting summary judgment only on headnotes whose language "very closely tracks the language of the Bulk Memo question but not the language of the case opinion," and found that the rest of the headnotes must go to trial. Order at 14. The court did not grant summary judgment to Ross on any headnotes because it did not believe that a reasonable jury could find any of them non-infringing. The court also left open for trial the factual question of which headnotes remained covered by Thomson Reuters's existing copyrights. The court next turned to Ross's defenses, rejecting all of them.

Innocent Infringement, Copyright Misuse, Merger, and Scenes a Faire Defenses

First, the court dismissed Ross's innocent infringement defense because that defense does not apply where an infringed work bears a copyright notice, which the Westlaw headnotes do. *See* 17 U.S.C. § 401(d). Second, the court dismissed the copyright misuse defense because, as it previously found in connection with dismissed antitrust counterclaims, Ross had not shown that Thomson Reuters misused its copyrights to "stifle competition." Third, the court found Ross's merger defense inapplicable because there are "many ways to express points of law from judicial opinions," such that Ross did not need to copy Thomson Reuters's protected expression. And fourth, the court dismissed the *scenes a faire* defense because that defense covers stock elements following from the nature of a work (*e.g.*, a damsel in distress in a romance novel) with no application on this set of facts.

Fair Use



court unplugs ai fair use defense, but context is key

The court left fair use for last. Previously, the court had denied summary judgment on the defense, but “with new information and understanding,” vacated those sections of its previous order and opinion. Noting that Ross bore the burden to prove its affirmative defense, the court engaged in the four-factor fair use analysis under Section 107 of the Copyright Act, balancing (1) the purpose and character of Ross’s use, including whether it is commercial; (2) the nature of the copyrighted work; (3) the amount and substantiality of the work used relative to the work’s whole; and (4) how Ross’s use affected the copyrighted work’s value or potential market.

Factor One. The court found that factor one weighed clearly in Thomson Reuters’s favor. Ross’s use was unquestionably commercial, which weighed against fair use. Citing the recent *Warhol* decision, the court found that Ross’s use was not transformative because it did not have a “further purpose or different character” from Thomson Reuters’s. *Andy Warhol Found. For the Visual Arts, Inc. v. Goldsmith*, 589 U.S. 508, 529-531 (2023). Ross used Westlaw headnotes to train a competing legal research tool, which ultimately functioned in a way similar to Westlaw’s headnotes and Key Number System—namely, to provide a user with “a list of cases with fitting headnotes.”

The court rejected Ross’s arguments that the use was transformative because the headnotes themselves do not appear in Ross’s final product and that its “intermediate copying” was transformative. Ross relied on cases that have found intermediate copies to be fair use involving software (particularly video games). *See, e.g., Sony Comput. Ent., Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000), *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992). Reversing its previous decision, the court found these cases to be inapt because they involved copying computer code (which did not happen in this case), and because the computer programs at issue almost always served functional purposes—meaning that, in the court’s view, the fair use considerations in intermediate copying cases do not always apply to cases involving the copying of written words. The intermediate copying cases also differed because in those cases, the copying was necessary for competitors to innovate, for instance so that programmers could allow different programs to speak to each other, or so that they could learn how to reverse engineer access to unprotected functional elements in computer programs. Here, the court stated, “there is no computer code whose underlying ideas can be reached only by copying their expression.” Order at 19. Looking at the “broad purpose and character” of Ross’s use in light of *Warhol*, the court found that Ross had simply copied the headnotes to make it easier to develop a competing product. For that reason, the court concluded the use was not transformative, but explicitly noted that “only non-generative AI” was before the Court.

Factor Two. The court found that this factor weighed in favor of Ross based on the minimal degree of creativity it found in the headnotes and Key Number System.

Factor Three. This factor also weighed in Ross’s favor because the output generated to an end user by Ross’s AI search tool did not include any Westlaw headnotes, and so Ross did not make headnotes available to the public. Order at 21.

Factor Four. The court found that Factor Four swung toward Thomson Reuters because Ross “meant to compete with Westlaw by developing a market substitute.” The court also found that Ross did not adduce sufficient evidence to show lack of a potential market for Thomson Reuters to sell its content as AI training data, nor evidence that such a market would be unaffected by Ross’s actions. Most interestingly, the court flatly rejected



court unplugs ai fair use defense, but context is key

Ross's arguments regarding any potential public interest that its product would serve. In the court's view, judicial opinions are freely available, and the public's "interest in the subject matter" alone is insufficient to find fair use. Order at 22-23, citing *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 569 (1985). Specifically, the court found that "[t]he public has no right to Thomson Reuters's parsing of the law," because "[c]opyrights encourage people to develop things that help society, like good legal-research tools," whose "builders earn the right to be paid accordingly." In short, there is "nothing that Thomson Reuters created that Ross could not have created for itself or hired LegalEase to create for it without infringing Thomson Reuters's copyrights." Order at 21-23.

Ultimately, the court rejected Ross's fair use defense because Factor One and Factor Four favored Thomson Reuters, and thus granted summary judgment for Thomson Reuters on fair use. Trial is currently set to begin on the remaining issues in this case on May 12, 2025.

Key Takeaways

The *Ross* case provides a window into how courts may deal with training issues in other AI litigation. It is nearly certain to be briefed in numerous other cases nationwide, likely very soon, and picked apart extensively. That said, it is a case with many specific quirks that could limit its applicability. For one, this is not a generative AI case. And as usual, the fair use finding is fact-specific and not broadly applicable to a multitude of other contexts. However, the principles discussed—such as the *Warhol* formulation of Factor One, the transformative nature of intermediate copying, and the Factor Four question of public interest—will be raised in other cases, and familiarity with this decision will be crucial going forward as the law develops.