Flattened fair-use factors fall flat with US judge at hearing on AI training data

6 Dec 2024 | 02:29 GMT | **Insight**
By Melissa Ritti

Counsel for Ross Intelligence and Thomson Reuters squared off for more than six hours in a US court today in a case with major ramifications for not only machine learning and artificial intelligence but also the legal research community as a whole. While the parties spent considerable time arguing the level of copyright protection the Westlaw database is entitled to and whether Ross's use of that same database was fair, the issue of substantial similarity — and a US judge's request for a side-by-side comparison of the works at issue — took center stage.

Counsel for Ross Intelligence and Thomson Reuters squared off for more than six hours in a US court today in a case with major ramifications for not only machine learning and artificial intelligence but also the legal research community as a whole.

While the parties spent considerable time arguing the level of copyright protection the Westlaw database is entitled to and whether Ross's use of that same database was fair, the issue of substantial similarity — and a US judge's request for a side-by-side comparison of the works at issue — took center stage.

US Judge Stephanos Bibas of the Third Circuit Court of Appeals, sitting by designation in the District of Delaware but today at the federal courthouse in nearby Philadelphia, presided over the six-hour summary judgment re-do in the high-stakes copyright clash between the two companies.

From whether "sweat of the brow," or a lack thereof, is an appropriate consideration when assessing copyright strength to accusations that Ross, not Thomson Reuters, is "killing the golden goose" of artificial intelligence by trying to compete with the Westlaw database, there was no shortage of metaphors — or areas of disagreement — at the hearing.

At issue are allegations of copyright infringement leveled over Ross's use of Westlaw to train its AI-powered legal research platform. It was assisted in that effort by LegalEase Solutions, whom Ross tasked with creating "memos" consisting of legal questions and answers.

LegalEase would go on to produce approximately 25,000 sets of data known as the "Bulk Memo Project." Thousands of those data sets, according to Thomson Reuters, were nothing more than Westlaw "headnotes" — a short summary of a longer passage in a judicial opinion, authored by Westlaw's attorney-editors. Ross, again via LegalEase, is also accused of making use of Westlaw's "key number system."

Headnotes link to the full-length judicial passages they correspond to, while key numbers direct users to a list of similar cases.

Like the vast majority of copyright cases involving AI training data playing out in US courts, the doctrine of fair use features heavily in Ross's defense. But unlike litigation by the *New York Times* or the Author's Guild, the case against Ross adds another dimension by presenting the question of what level of protection the Westlaw elements — registered as a "compilation of legal material" — are entitled to, if any at all.

Separately, the dispute requires a factfinder to assess how similar or dissimilar the headnotes are to the underlying, corresponding judicial opinions, which are not copyrightable, and lastly, whether they are similar or dissimilar to the bulk memo data sets.

Bibas denied competing motions for summary judgment last year. However, in August 2024, he invited the parties to submit two new sets of briefings on copyrightability, validity and infringement, as well as Ross's fair use defense. The parties filed their renewed summary judgment briefs beginning in early October (see here).

Late last month, in a ruling from the bench, Bibas directed the parties to be prepared to argue seven issues at today's hearing: fair use factors one and four, the type of copyright being claimed by Reuters in its headnotes and key number system, whether the headnotes should be evaluated on their own or as a complete set, whether substantial similarity is a question for the judge or a jury, and "the relationship, if any, between Fed. R. Civ. P. 54(b) and the law of the case

doctrine, as the Court is considering the extent to rely upon or revisit its prior summary judgment opinion."

At times, the judge appeared surprised that after four years of litigation he had not yet been presented with a side-by-side comparison of the bulk memo questions and answers, the headnotes from Westlaw they allegedly copied, and the corresponding judicial opinion from which each Westlaw headnote was derived.

Warrington Parker, counsel for Ross, vowed to provide as much, agreeing with Bibas that if the case advances to trial on damages the number of headnotes infringed will become relevant. "Right now, they're asking us to disgorge everything," Parker said.

Joshua Simmons, representing Thomson Reuters, was also amenable to presenting Bibas with a side-by-side comparison. "Ross's argument is that (the headnotes) are too close to the judicial opinions. The court can easily decide that for himself," Simmons said.

During the hearing Bibas made apparent his strong preference for Third Circuit and Supreme Court precedent in addressing Parker's claim that the more transformative the use at fair use factor one, the less likely it is to cause harm at factor four. "It sounds like you're saying factor one is a threshold issue," Bibas said.

Parker demurred, telling the court he did not mean to conflate the two, but later reiterated they are "related" under Sega Enterprises Ltd. v. Accolade and Sony Computer Entertainment v. Connectix — both decided by the Ninth Circuit.

"They don't bind me," Bibas said. "Why should they persuade me?"

Dale Cendali, representing Thomson Reuters, focused heavily on Ross's goal of siphoning off business from Westlaw, and the US Supreme Court ruling last year in Andy Warhol Foundation for the Visual Arts v. Goldsmith that commercial intent is relevant to the fair use analysis and that the fair use factors are non-exhaustive.

"Any way you slice it, this was harmful," Cendali said.

The hearing was expected to last two days but concluded this evening. Trial in early 2025 is possible, Bibas previously indicated.

Please e-mail editors@mlex.com to contact the editorial staff regarding this story, or to submit the names of lawyers and advisers.

Related Portfolio(s):
Artificial Intelligence - Antitrust - Thomson Reuters - ROSS Intelligence - Copyright litigation over Ross use of Westlaw to develop AI legal chatbot (US)

Industries: Computing & Information Technology, Consumer Products, Professional Services

Geographies: Delaware, North America, Northeast USA, United States

Topics:

Artificial Intelligence (AI) Copyright