

No.

**In The United States Court of Appeals
For the Third Circuit**

THOMSON REUTERS ENTERPRISE CENTRE GMBH and WEST
PUBLISHING CORPORATION,
Plaintiffs-Respondents,

v.

ROSS INTELLIGENCE INC.,
Defendant-Petitioner.

*On Petition for Certification from the
United States District Court for the District of Delaware
Civil Action No. 20-613 (The Honorable Stephanos Bibas)*

PETITION FOR CERTIFICATION UNDER 28 U.S.C. § 1292(b)

Anne M. Voigts
Ranjini Acharya
PILLSBURY WINTHROP
SHAW PITTMAN LLP
2550 Hanover Street
Palo Alto, CA 94304

Yar R. Chaikovsky
Andy LeGolvan
WHITE & CASE LLP
3000 El Camino Real
2 Palo Alto Square
Suite 900
Palo Alto, CA 94306

Mark S. Davies
Counsel of Record
Anna B. Naydonov
Kufere J. Laing
WHITE & CASE LLP
701 Thirteenth Street, NW
Washington, DC 20005
Tel: (202) 626-3600
mark.davies@whitecase.com

Kayvan Ghaffari
PILLSBURY WINTHROP
SHAW PITTMAN LLP
Four Embarcadero
Center, 2nd Floor
San Francisco, CA 94111

Warrington S. Parker III
Jacob Canter
CROWELL & MORING LLP
3 Embarcadero Ctr.,
San Francisco, CA 94111

Counsel for Defendant-Petitioner ROSS Intelligence, Inc.

CORPORATE DISCLOSURE STATEMENT

Pursuant to Rule 26.1 of the Federal Rules of Appellate Procedure and Third Circuit Local Appellate Rule 26.1(b), Petitioner ROSS Intelligence, Inc. discloses that it has no parent corporations, does not issue stock, and that there is no publicly held corporation which is not a party to the proceeding before this Court with a financial interest in the outcome of the proceeding.

TABLE OF CONTENTS

INTRODUCTION..... 1

STATEMENT OF THE CASE 2

 I. Founding and Rapid Growth of ROSS Intelligence, Inc. 2

 II. Certified Questions 6

RELIEF SOUGHT..... 8

REASONS TO GRANT PERMISSION TO APPEAL..... 8

 I. The originality and fair use questions present controlling
 questions of law..... 9

 II. There are substantial grounds for differences of opinions
 on the originality and fair use questions..... 10

 A. There are substantial grounds for a difference of
 opinion on the originality of Westlaw headnotes. 11

 B. There are substantial grounds for a difference of
 opinion on ROSS’s use of Westlaw headnotes..... 15

 III. An interlocutory appeal will materially advance this
 litigation. 23

 IV. Resolving these critically important questions would serve
 the public interest. 25

CONCLUSION..... 26

TABLE OF AUTHORITIES

Page(s)

CASES

Ahrenholz v. Bd. of Trustees of Univ. of Ill.,
 219 F.3d 674 (7th Cir. 2000) 9, 10

Am. Geophysical Union v. Texaco Inc.,
 802 F. Supp. 1 (S.D.N.Y. 1992)..... 15

Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith,
 598 U.S. 508 (2023)..... 16, 18

Authors Guild v. Google, Inc.,
 804 F.3d 202 (2d Cir. 2015) 20, 21, 22

Banks v. Manchester,
 128 U.S. 244 (1888)..... 13

Campbell v. Acuff-Rose Music, Inc.,
 510 U.S. 569 (1994)..... 16, 21, 22

Coopers v. Lybrand & Livesay,
 437 U.S. 463 (1978)..... 8

Feist Publ’n, Inc. v. Rural Tel. Serv. Co.,
 499 U.S. 340 (1991)..... 9, 11

Google v. Oracle,
 593 U.S. 1 (2021)..... *passim*

Katz v. Carte Blanche Corp.,
 496 F.2d 747 (3d Cir. 1974) 9, 10, 23, 24

Knipe v. SmithKline Beecham,
 583 F. Supp. 2d 553 (E.D. Pa. 2008)..... 24

Link v. Mercedes-Benz,
 550 F.2d 860 (3d Cir. 1977) 9, 10

Matthew Bender & Co. v. West Publ’g Co.,
 158 F.3d 674 (2d Cir. 1998) 13, 14

Rees v. BP Expl. (Alaska) Inc.,
 643 F.3d 681 (9th Cir. 2011)..... 10

Stewart v. Abend,
 495 U.S. 207 (1990)..... 16

Worth v. Selchow & Righter Co.,
 827 F.2d 569 (9th Cir. 1987)..... 13

STATUTES AND RULES

17 U.S.C. § 102 1

17 U.S.C. § 107 1, 15, 16

28 U.S.C. § 1292(b)..... 1, 8, 23, 24

MISCELLANEOUS

90 Fed. Reg. 9088 (Feb. 6, 2025) 25

Andrew Arruda, *The world’s first AI legal assistant*, YouTube (Dec. 21,
 2016), <https://tinyurl.com/nzbnkm7hb> 3

16 Charles A. Wright & Arthur R. Miller, *Federal Practice and
 Procedure* § 3929 Permissive Interlocutory Appeals-Section 1292(b)
 (3d ed.)..... 23

Chat GPT is Eating the World, *Updated Map of 40 US copyright suits v.
 AI companies, with transfer of all suits v. OpenAI to Judge Stein*,
<https://tinyurl.com/Allawsuits>..... 25

Christopher Lehane, Public Comment on behalf of OpenAI,
<https://tinyurl.com/58m6cfwx> 26

Harlan F. Stone, *The Common Law in the United States*, 50 Harv. L.
 Rev. 4 (1936) 25

Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105 (1990).... 21

Statement of the Honorable John J. Parker, Legislative History of the
Judicial Code Amendments, (1958)..... 23, 24

INTRODUCTION

More than 10 years ago, the founders of ROSS Intelligence, Inc. started work on an AI search engine. To “train” the engine, ROSS applied advanced mathematical and statistical algorithms (explained more below) to .076% of the over 28 million headnotes available on westlaw.com. The engine’s answers were quotes from relevant judicial opinions, and the answers did not include any Westlaw content. As the market started to recognize ROSS’s value, Thomson Reuters filed this copyright suit.

After first deciding that questions of originality (17 U.S.C. § 102) and “fair use” (*id.* § 107) had to be decided by the jury, the district court (Bibas, J.) reversed course days before trial, and incorrectly held that—as a matter of law—the West headnotes were original and ROSS’s use of those headnotes was not fair. Recognizing, however, that there were “substantial grounds for a difference of opinion on the controlling legal issues in this case” that “have the potential to change the shape of trial,” the district court then certified two questions for interlocutory review under 28 U.S.C. § 1292(b):

- 1) whether the West headnotes and Key Number System are original; and

2) whether ROSS's use of the headnotes was fair. App. 1.

This Court should grant ROSS's petition for interlocutory appeal to address the important copyright issues presented.

STATEMENT OF THE CASE

I. Founding and Rapid Growth of ROSS Intelligence, Inc.

In 1887, the American Digest System provided a method for indexing law. This method was based on preexisting casebooks and digests. Just over two decades later, West Publishing Corporation introduced the West Key Number System. Over 75% of the topics in Westlaw's system are the same topics as in the American Digest. The points of law are included in "headnotes" with "key numbers." The first rule for creating those headnotes is "follow[ing] the court's language" "insofar as possible" to achieve "[a]ccuracy." *See* D.E. 787, ROSS Opening Brief in Support of Interlocutory Appeal ("ROSS IA"), at *6. Westlaw made its collection of public caselaw with headnotes available in books, and, beginning in 1998, on westlaw.com. *See* D.E. 677, Redacted Renewed ROSS Opening Brief in Support of Summary Judgment on Fair Use ("ROSS OB") at *11–*12. By 2010, Westlaw had developed a search algorithm to find headnotes, judicial opinions, and other sources. Today,

Westlaw's system includes over 28 million headnotes and more than 100,000 points of law in 14 million judicial opinions. Ross OB at *30–*31.

Given this trove of public judicial opinions and the difficulty of accessing all this law, in 2014, Andrew Arruda, Jimoh Ovbiagele, and Pargles Dall'Oglio entered a student competition at the University of Toronto sponsored by IBM. Using Watson, IBM's cognitive computing foundational model, the student-competitors were asked to build something. Arruda, Ovbiagele, and Dall'Oglio set out to design artificial intelligence to improve searching. They won the competition, and ROSS Intelligence, Inc. ("ROSS") was born. A year later, ROSS was accepted into Y Combinator for further growth. Its meteoric rise continued. *See* ROSS OB at *4.

ROSS had one goal: to reduce the cost of legal research and increase the accessibility of the law by directing the public straight to the language contained in judicial opinions—not to lists of cases or secondary sources. *See* Andrew Arruda, *The world's first AI legal assistant*, YouTube (Dec. 21, 2016), <https://tinyurl.com/nzbkm7hb>. ROSS had access to a repository of judicial opinions through a company named Casemaker. But ROSS needed to train a domain-aware AI system to

select the relevant paragraphs from judicial opinions that responded to a user's question.

No small task. AI systems respond to math; but humans write judicial opinions and ask questions in prose. ROSS needed to build an information retrieval algorithm that could bridge that gap. Said differently, ROSS's algorithm needed to (1) use math to understand the questions and (2) use the judicial opinions to answer the questions accurately.

To that end, ROSS used memos that paired a tiny percentage (.076%) of headnotes available on Westlaw as questions with answers excerpted from judicial opinions. *E.g.*, App. 39. There were 25,000 question-answer pairs in total. ROSS OB at *3. The question-and-answer pairs were, then, placed into .csv files. Each line of the .csv file contained a question, a judicial passage as an answer, and a relevance rating that was associated with the question-answer pair. From there, the text was prepared for computational analysis and ROSS applied various mathematical and statistical algorithms to generate numerical "features," which quantified the similarity between specific attributes of the question-and-answer texts. D.E. 680-1, Ovbiagele Decl. ¶¶ 15–19.

Each question-answer pair was represented by more than two dozen features. Once armed with those features, ROSS no longer needed, nor used, the question-answer pairs. Rather, ROSS turned to a machine-learning tool that combined and generated hundreds of thousands of composite features to train its search model. As training continued, the model learned to adjust the features' weights based on the ratings in the training data to create a model. Ultimately, the model learned to rank how effectively caselaw passages answer a question. D.E. 680-1, Ovbiagele Decl. ¶¶ 21–28.

The result was ROSS's AI search engine. A user asked a question. The AI system, trained to take the question, searched across the millions of judicial opinions (purchased from Casemaker), and then provided the user with a list of passages from relevant judicial opinions. The user was not provided with any Westlaw headnotes or other Westlaw content. The ROSS AI search engine did not use the Key Number System in any way. And each use of the search system taught the engine additional relevant information about the dataset. The initial training became less important over time.

II. Certified Questions

In 2020, after ROSS was attracting an ever-increasing number of monthly subscribers and investor interest, Thomson Reuters, West Publishing's owner, sued ROSS alleging copyright infringement and tortious interference with contract. The lawsuit forced ROSS to quickly shut down its operations. Hoping to reopen its doors, ROSS counterclaimed requesting declaratory judgments on copyright invalidity, non-infringement, and no tortious interference with contract.

On April 29, 2022, as the parties were engaged in discovery, Judge Bibas became the presiding judge following Judge Stark's confirmation to the United States Court of Appeals for the Federal Circuit. *See* Joint Status Report, D.E. 180. Six months later, in December 2022, the parties began the first round of summary judgment briefing on the copyright, copyright preemption, and tortious interference with contract claims. Judge Bibas issued *Thomson Reuters I* on September 27, 2023, finding genuine disputes of material fact on copyright infringement and fair use. App. 2. A little less than a month later, Judge Bibas informed the parties that a trial on copyright issues would begin on August 26, 2024. D.E. 581.

But just days before trial, on August 22, 2024, Judge Bibas continued the trial sua sponte without setting a future date. D.E. 663. He then invited the parties to renew their summary judgment briefing on copyrightability, validity, infringement, and fair use. Judge Bibas heard oral argument on these motions in December 2024, and published *Thomson Reuters II* on February 11, 2025. App. 36. In so doing, Judge Bibas concluded, among other things, that he could decide originality and fair use as a matter of law—and find in Thomson Reuters’s favor.

A little over a month later, ROSS moved for interlocutory appeal under 28 U.S.C. § 1292(b) and to stay the trial pending this Court’s review. Though “confident” in the “February 2025 summary judgment opinion,” Judge Bibas recognized that there “are substantial grounds for a difference of opinion on controlling legal issues in this case.” App. 1. Because these “issues have the potential to change the shape of trial,” Judge Bibas promptly granted ROSS’s motion on April 4, 2025, days after Thomson Reuters filed its brief in opposition, and without ROSS filing a reply brief. *Id.*

ROSS now petitions this Court to answer two questions:

1. Do the headnotes fail the Copyright Act's originality requirement because the notes lack "creative spark"?
2. Was ROSS's use of the headnotes transformative, or otherwise fair?

RELIEF SOUGHT

ROSS requests that this Court grant this petition for permission to appeal under 28 U.S.C. § 1292(b) and, on appeal, hold that (i) Westlaw headnotes are unoriginal, and, in any event (ii) ROSS's use is fair.

REASONS TO GRANT PERMISSION TO APPEAL

This is an appropriate case for interlocutory review. Congress enacted Section 1292(b) as an exception to the final judgment rule "to meet the recognized need for prompt review of certain nonfinal orders." *Coopers v. Lybrand & Livesay*, 437 U.S. 463, 474 (1978). Under Section 1292(b), appellate courts have jurisdiction over interlocutory orders that: (1) involve a controlling question of law as to which (2) there is a substantial ground for a difference of opinion and that (3) an immediate appeal from that order may materially advance the litigation. 28 U.S.C. § 1292(b).

Although the district court got the underlying order wrong, it was right in recognizing that this case satisfies all three requirements on each question. And though this Court’s signoff is necessary to protect its docket from becoming overcrowded with piecemeal or minor litigation, the certified questions are of unique importance that raise issues of fundamental importance about artificial intelligence, copyright, and national security that are likely to arise time and again. This case deserves the Court’s immediate attention.

I. The originality and fair use questions present controlling questions of law.

“A controlling question of law must encompass at the very least every order which, if erroneous, would be reversible error on final appeal.” *Katz v. Carte Blanche Corp.*, 496 F.2d 747, 755 (3d Cir. 1974). “These orders [are] definitive, effective, and in a posture capable of affirmance and reversal.” *Link v. Mercedes-Benz*, 550 F.2d 860, 863 (3d Cir. 1977). In short, a “controlling” question is “a pure question of law, something the court of appeals [can] decide quickly and cleanly without having to study the record.” *Ahrenholz v. Bd. of Trustees of Univ. of Ill.*, 219 F.3d 674, 676 (7th Cir. 2000) (Posner, J.).

“Copyright validity is a question of law, not fact.” App. 40. In

particular, as relevant here, originality is the “touchstone of copyright protection.” *Feist Publ’n, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 347 (1991). The first certified question, whether Westlaw headnotes are original, is thus a pure question of law that should cleanly resolve this case.

Fair use as presented here is also a question of law. As explained in *Thomson Reuters II*, sometimes fair use is “a mixed question of law in fact,” but not here—the “undisputed facts ... push this case squarely into the legal realm.” App. 51. ROSS agrees. Its use of the headnotes is fair as a matter of law. There is no need for the Court to “hunt[] through the record compiled in the summary judgment proceeding to see whether there may be a genuine issue of material fact lurking there.” *Ahrenholz*, 219 F.3d at 677. In short, the relevant orders that give rise to this appeal are “in a posture capable of ... reversal.” *Link*, 550 F.2d at 863.

II. There are substantial grounds for differences of opinions on the originality and fair use questions.

When “novel legal issues are presented, on which fair-minded jurists might reach contradictory conclusions,” there are substantial grounds for a difference of opinion that warrant interlocutory review. *Rees v. BP Expl. (Alaska) Inc.*, 643 F.3d 681, 688 (9th Cir. 2011). And

there “can be little difficulty over the criterion of difference of opinion,” where, as here, “it took two district court opinions to arrive at a decision.” *Katz*, 496 F.2d at 755.

A. There are substantial grounds for a difference of opinion on the originality of Westlaw headnotes.

1. “To qualify for copyright protection, a work must be original to the author.” *Feist*, 499 U.S. at 345. Originality follows a simple formula: “independent creation plus a modicum of creativity.” *Id.* at 346. And it is a “bedrock principle of copyright that ... [n]o one may claim originality as to facts.” *Id.* at 347 (internal quotation marks and citations omitted). “Factual compilations, on the other hand, may possess the requisite originality” depending on the “selection” and “arrangement” of the facts. *Id.* at 348.

But this protection has a clear limiting principle: “copyright protection may extend only to those components” of a compilation that “are original to the author.” *Id.* That is why “copyright in a factual compilation is thin,” especially where “the compilation author adds no written expression but rather lets the facts speak for themselves.” *Id.* Plain and simple, “facts themselves do not become original through association.” *Id.*

Judge Bibas’s summary judgment orders reach different holdings regarding the originality of the headnotes. *Compare* App. 8–9 *with* App. 42–43. The dueling opinions reflect the substantial grounds for a difference of opinion on originality.

Thomson Reuters I declared that the headnotes “are not aptly described by the compilation caselaw” and must be considered “at the level of each individual headnote.” App. 8. “If a headnote merely copies a judicial opinion, it is uncopyrightable,” and ROSS “presents evidence that Thomson Reuter’s protocols required headnotes to follow or closely mirror the language of judicial opinions.” *Id.* Accordingly, *Thomson Reuters I* concluded that there was a “genuine factual dispute about how original [they] are,” which “affects the strength and extent of Thomson Reuters’s copyright, and ... whether ROSS was copying the headnotes or the opinions themselves.” App. 9.

Thomson Reuters II jettisoned *Thomson Reuters I*’s reasoning and its conclusions. “First,” it found that “headnotes are a compilation” and therefore “Thomson Reuters’s selection and arrangement of its headnotes easily” clear *Feist*’s “minimal degree of creativity.” App. 42. “More than that,” *Thomson Reuters II* continued, “each headnote is an individual

copyrightable work” even those that track a judicial opinion “verbatim.” *Id.* That is because each headnote is “a carefully chosen fraction of the whole, identifying which words matter” from the judicial opinion, the district court further reasoned. *Id.* The “editorial expression” in selecting the relevant judicial language, *Thomson Reuters II* concluded, “has enough ‘creative spark’ to be original.” App. 41–42.

2. The substantial grounds for a difference of opinion in *Thomson Reuters I* and *II* are only the start. Both opinions are wrong. Westlaw headnotes, which parrot judicial opinions as best they can, do not contain a modicum of creativity. Not individually. Nor as a compilation. Facts are not copyrightable. Judicial opinions are not either. *See Banks v. Manchester*, 128 U.S. 244 (1888). That is a matter of law. *See, e.g., Worth v. Selchow & Righter Co.*, 827 F.2d 569, 572–73 (9th Cir. 1987) (noting that the facts contained within a compilation are not protected under the Copyright Act).

Just ask the Second Circuit. In *Matthew Bender & Co.*, it concluded that “West’s alterations to [its] case reports, when considered collectively” failed to “demonstrate sufficient originality and creativity to be copyrightable.” *Matthew Bender & Co. v. West Publ’g Co.*, 158 F.3d

674, 680, 689 (2d Cir. 1998). “The originality standard requires” a “creative spark,” and that spark “is missing,” the court explained, “where: (i) industry conventions or other external factors so dictate selection that any person composing a compilation of the type at issue would necessarily select the same categories of information ... or (ii) the author made obvious, garden-variety, or routine selections.” *Id.* at 682. The “selection or arrangement of information” in Westlaw’s case reports lacked the “creativity inhere[d] in making non-obvious choices from among more than a few options.” *Id.*

The headnotes share those same unoriginal defects. Indeed, “industry conventions” and functionality dictate the Westlaw headnote editors’ decisions. *Id.* at 682. The editors are instructed to “follow the court’s language” “insofar as possible” to achieve “[a]ccuracy.” ROSS IA at *6. And copyright law does not extend to recitation of fact. So it is legal error to conclude that “a headnote taken verbatim from an opinion” is original because it “is a carefully chosen fraction of the whole.” App. 42. As *Matthew Bender & Co.* put it, “for West or any other editor of judicial opinions for legal research, faithfulness to the public-domain original is the dominant editorial value, so that the creative is the enemy of the

true.” 158 F.3d at 688.

B. There are substantial grounds for a difference of opinion on ROSS’s use of Westlaw headnotes.

“Generally speaking, the decisions of the Supreme Court on fair use have not formulated a clear framework or standard governing future cases.” *Am. Geophysical Union v. Texaco Inc.*, 802 F. Supp. 1, 30 (S.D.N.Y. 1992) (Leval, J.) (certifying the fair use question after finding it inapplicable). The doctrine’s ambiguities that Judge Leval observed in 1992 are a doctrinal feature, not a temporal bug. That is because fair use will always reflect its judge-made origins that Congress has codified into statute.

As the Supreme Court recently explained, the statute’s enumerated, inexhaustive, factors: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work, 17 U.S.C. § 107, require “judicial balancing, depending upon relevant circumstances, including significant changes in technology,” *Google v. Oracle*, 593 U.S. 1, 19 (2021) (cleaned up).

Given the doctrine’s “equitable” roots, the Supreme Court has repeatedly cautioned courts “to avoid rigid application of the copyright statute when ... it would stifle the very creativity which that law is designed to foster.” *Id.* at 18 (quoting *Stewart v. Abend*, 495 U.S. 207, 236 (1990)). Fair use is “clear” on only one point—the “concept is flexible,” “courts must apply it in light of the sometimes conflicting aims of copyright law, and [] its application may very well vary depending upon context.” *Id.* at 20.

The district court’s two conflicting opinions reflect substantial disagreement on the application and balancing of the fair use factors.

The purpose and character of ROSS’s use. The first fair use factor “considers the reasons for, and nature of, the copier’s use of an original work.” *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528 (2023). Research, as “one of the purposes listed in the preamble paragraph of Section 107,” illustrates the sort “of copying that courts and Congress most commonly have found to be fair uses,” and thus “guide[s] the first factor inquiry.” *Id.* (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 (1994)). At bottom, this factor seeks to determine whether the “new work” “adds something new, with a further

purpose or different character,” such that it is “transformative.” *Campbell*, 510 U.S. at 579.

Thomson Reuters I and *Thomson Reuters II* vehemently disagree over whether ROSS’s use is transformative. *Thomson Reuters I* credited ROSS’s use of AI to translate “human language into something understandable by a computer as a step in the process of trying to develop a ‘wholly new,’ albeit competing, product—a search tool that would produce highly relevant quotations from judicial opinions in response to natural language questions.” App. 20. Viewed this way, *Thomson Reuters I* concluded that a jury must decide whether ROSS engaged in “transformative intermediate copying.” *Id.*

Thomson Reuters II disclaimed any possibility of transformative intermediate copying. The intermediate copying caselaw, it asserted, protects “copying computer code,” not prose. App. 53. And in those cases, the opinion continued, copying code “was necessary” for innovation. *Id.* Because ROSS did not copy computer code, and copying headnotes were not necessary to build ROSS’s AI search, *Thomson Reuters II* found that ROSS’s use of numerical code to find “relevant judicial opinions” “resembles how Westlaw uses headnotes and key numbers.” App. 52.

Thus, all that mattered was that ROSS “took the headnotes to make it easier to develop a competing legal research tool,” so its use was commercial, “not transformative.” App. 54.

Thomson Reuters I and *II*'s disagreement aside, both opinions are doctrinally flawed. The “central question” factor one asks is whether the “new work” “adds something new, with a further purpose or different character.” *Goldsmith*, 598 U.S. at 528 (internal quotation marks and citations omitted). Yet neither opinion adequately engages with, let alone answers that question here, where ROSS built a search tool that uses AI to match natural language questions with excerpts straight from judicial opinions. Surely, the use of AI to assist with legal research is of a different character than forcing users to manually sort through scores of headnotes.

The effect of the use upon the market for the work. *Thomson Reuters I* and *II* are irreconcilable here too. This factor may require consideration of “the amount of money that the copyright owner might lose,” “the source of the loss,” and “the public benefits the copying will likely produce.” *Google*, 593 U.S. at 35. Deducing that ROSS’s “use might be transformative,” *Thomson Reuters I* proffered, ROSS’s “brand-new

research platform” may serve “a different purpose than Westlaw. If so, it is not a market substitute.” App. 26. *Thomson Reuters I* also credited ROSS’s argument that “Thomson Reuters has never participated—and would never participate—in” a market for its headnotes as training data. *Id.* As for the public benefit, *Thomson Reuters I* articulated a “hotly debated question,” that it concluded only a jury could answer: “Is it in the public benefit to allow AI to be trained with copyrighted material?” *Id.*

Thomson Reuters II discarded *Thomson Reuters I* as “unpersuasive.” App. 57. In its view, all that mattered was that ROSS “meant to compete with Westlaw by developing a market substitute” and it failed to “put forward enough facts to show” that the “potential market for AI training data” does not “exist and would not be affected.” *Id.* Whatever public benefit ROSS offers, *Thomson Reuters II* continued, is negligible because the “public has no right to Thomson Reuters’s parsing of the law.” App. 58.

Errors abound. The district court’s theory is as unprecedented as it is wrong. Neither *Thomson Reuters I* nor *II* cite a single case that supports the proposition that factor four considers the potential effect on a hypothetical market. Under the district court’s logic, copyright law

imposes another burden on innovators—to affirmatively show that they did not create a new-hypothetical market; and, in the alternative, prove that this hypothetical market did not affect an alleged competitor. That is stunning.

Thomson Reuters II's discussion of the public benefit is equally jaw dropping. It boldly asserted that the “public has no right to Thomson Reuters’s parsing of the law.” App. 58. That assertion is irrelevant to this case. To reiterate, ROSS did not give the public Thomson Reuters’s parsing of the law—it provided its users with verbatim quotes from uncopyrightable judicial opinions devoid of any third-party content. And there can be no dispute, that increasing access to the law with “good legal-research tools” helps society. App. 58.

Undeterred, *Thomson Reuters II* attempted to distinguish this case “from *Google*, where the API was valuable ‘because users, including programmers, were just used to it.’” *Id.* (quoting *Google*, 593 U.S. at 38). But this case is just like *Google*—Westlaw has been on the market for over a century, and has become valuable to lawyers for the exact reason the API was valuable for developers “users ... are just used to it. They have already learned to work with it.” *Google*, 593 U.S. at 38–39. There

is “no reason to believe that the Copyright Act seeks to protect third parties’ investment in learning how to operate a created work.” *Id.* at 39.

Balancing. *Thomson Reuters I* and *II* share a controversial premise: the “first and fourth factors weigh most heavily in the analysis.” App. 51 (citing *Authors Guild v. Google, Inc.*, 804 F.3d 202, 220 (2d Cir. 2015) (Leval, J.)); *see also* App. 15. But *Authors Guild* says no such thing. And this Court has not hinted at—let alone endorsed—such a per se rule.

Nor has the Supreme Court. All four factors “are to be explored, and the results weighed together, in light of the purposes of copyright law.” *Campbell*, 510 U.S. at 578. “The task is not to be simplified with bright-line rules, for the statute, like the doctrine it recognizes calls for case-by-case analysis.” *Id.* at 577. Though “some factors may prove more important in some contexts than in others,” *Thomson Reuters I & II*’s bright-line rule mandates the very thing that fair use rejects—a “rigid application” that ignores “relevant circumstances.” The district court betrayed these first principles and transformed fair use balancing into its antithesis, a “score card that promises victory.” *Google*, 593 U.S. at 19 (quoting Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105, 1110 (1990)). *Thomson Reuters I* and *II* begin on the wrong foot (and thus

end on it too).

In turn, neither decision adequately engages a critical, and undisputed, fact: ROSS used .076% of Westlaw headnotes. In *Google*, for example, the “quantitative amount copied” (.4%) and its “basic purpose” (“programmers had already learned to work” with the system) played a central role in the fair use analysis. *Google*, 593 U.S. at 33, 34. As *Authors Guild* explained, the third factor’s “clear implication” is that “a finding of fair use is more likely when small amounts, or less important passages, are copied.” 804 F.3d at 221. “The obvious reason for this lies in the relationship between the third and fourth factors.” *Id.* That is a far cry from *Thomson Reuters I* and *II*’s per se rule that cites *Authors Guild* for support.

Indeed, *Thomson Reuters I* and *II* buck all of this. In the latter opinion, the district court deemphasized the fact that ROSS used “only a small percentage of the total headnotes owned by Westlaw” and instead focused on the fact that ROSS “did not make West headnotes available to the public.” App. 56. Though the district court ultimately concluded the “amount and substantiality of the portion used” favored ROSS’s fair use, the disregard of ROSS’s *de minimis* copying significantly downplayed the

weight courts afford to users who barely copy the protected work and threw the court's balancing into disarray. *Id.*; see also *Campbell*, 510 U.S. at 578 (all four fair use factors “are to be explored, and the results weighed together, in light of the purposes of copyright”).

III. An interlocutory appeal will materially advance this litigation.

Section 1292(b) must be applied consistently with the “policies favoring interlocutory appeal,” one of which is “the avoidance of possibly wasted trial time and litigation expense.” *Katz*, 496 F.2d at 756. As “the person most familiar with the litigation,” the trial court’s determination that an order is ripe for appellate review is entitled to deference. 16 Charles A. Wright & Arthur R. Miller, *Federal Practice and Procedure* § 3929 Permissive Interlocutory Appeals—Section 1292(b) (3d ed.). Here, as the district court explained in its initial order granting ROSS’s motion for interlocutory appeal, the copyright “issues have the potential to change the shape of the trial.” App. 1.

That conclusion is consistent with the “typical example” that satisfies Section 1292’s “materially advance” element: eliminating the possibility that a “circuit court of appeals” will reverse a district court’s determination that a theory of legal liability or an affirmative defense is

“not applicable” and goes on through a long trial that results in damages. Statement of the Honorable John J. Parker, Legislative History of the Judicial Code Amendments, 1958: P.L. 85-554; and, P.L. 85-791; and, P.L. 85-919; and, P.L. 85-920, at *8 (1958) (“Statement of Judge Parker”). If the court of appeals reverses the trial court’s ruling on liability or the applicability of the affirmative defense, it must also vacate the damages that the jury awarded, and ultimately the district court’s “work [was] for nothing.” Statement of Judge Parker at *8; see also *Katz*, 496 F.2d at 754.

Immediate appellate review here would “materially advance the litigation.” 28 U.S.C. § 1292(b). To make things plain, *Thomson Reuters II* asked a jury to resolve myriad copyright issues that are necessary only if: (1) Westlaw’s headnotes are original *and* (2) ROSS’s use of the headnotes is unfair. But if this Court vacates and reverses on either ground, then the trial is for naught: jury selection, listening to days of testimony, hours (maybe days) of deliberation, and calculating damages—efforts that are wasted. The district court certified these questions rather than force jurors take on these important tasks, uncertain of their necessity. This Court’s resolution of the novel copyright issues, thus, materially advances the litigation.

The fact that the tortious interference with contract claim remains does not counsel against immediate appellate review. If this Court vacates and remands on the copyright claims, a one-issue tortious interference trial is far simpler than a two-issue trial involving novel copyright issues. *See Knipe v. SmithKline Beecham*, 583 F. Supp. 2d 553, 600 (E.D. Pa. 2008) (immediate review will materially advance the ultimate termination of litigation if “the trial would be simplified by the elimination of complex issues.”).

IV. Resolving these critically important questions would serve the public interest.

The question of how copyright law applies to artificial intelligence innovations is one of the most, if not the most, fundamental legal question of today.¹ Copyright law, and fair use especially, is a common law doctrine that “performs its function adequately only when it is suited to the way of life of people.” Harlan F. Stone, *The Common Law in the United States*, 50 Harv. L. Rev. 4, 11 (1936).

The relationship between AI and national security currently

¹ See Chat GPT is Eating the World, *Updated Map of 40 US copyright suits v. AI companies, with transfer of all suits v. OpenAI to Judge Stein*, <https://tinyurl.com/AIlawsuits>.

commands the Executive Branch’s attention. That is why it has recently invited public comment on the development of an Artificial Intelligence Action Plan. *See* 90 Fed. Reg. 9088 (Feb. 6, 2025). Some commenters have shown that applying the “fair use doctrine to AI” is a “matter of national security.” *See, e.g.*, Christopher Lehane, Public Comment on behalf of OpenAI, at *10, <https://tinyurl.com/58m6cfwx>. That is because foreign corporations that enjoy “concerted state support” have “unfettered” access to “copyrighted data” that will “improve” their AI models. *Id.* But under the district court’s opinion, “American companies are left without fair use access,” and the “race for AI is effectively over.” *Id.*

CONCLUSION

The Court should grant the petition for interlocutory appeal.

Respectfully submitted,

/s/Mark S. Davies

Mark S. Davies

Counsel of Record

Yar R. Chaikovsky
Andy LeGolvan
WHITE & CASE LLP
3000 El Camino Real
2 Palo Alto Square
Suite 900
Palo Alto, CA 94306

Anna B. Naydonov
Kufere J. Laing
WHITE & CASE LLP
701 Thirteenth Street, NW
Washington, DC 20005
Tel: (202) 626-3600

Warrington S. Parker III
Jacob Canter
CROWELL & MORING LLP
3 Embarcadero Ctr., 26th Floor
San Francisco, CA 94111

Anne M. Voigts
Ranjini Acharya
PILLSBURY WINTHROP SHAW
PITTMAN LLP
2550 Hanover Street
Palo Alto, CA 94304

Dated: April 14, 2025

*Attorneys for Defendant-Petitioner
ROSS Intelligence, Inc.*

Kayvan Ghaffari
PILLSBURY WINTHROP SHAW
PITTMAN LLP
Four Embarcadero Center, 2nd
Floor
San Francisco, CA 94111

CERTIFICATE OF COMPLIANCE

I hereby certify that (i) this document complies with the word limit of Federal Rule of Appellate Procedure 5(c)(1) because it contains 5,077 words exclusive of the portions exempted by Federal Rule of Appellate Procedure 32(f); (ii) this document complies with the typeface requirement of Federal Rule of Appellate Procedure 32(a)(5)(A) because it was typed in 14-point Century Schoolbook font; (iii) I am a member in good standing of the bar of the United States Court of Appeals for the Third Circuit; and (iv) the electronic copy of this brief has been virus scanned using Microsoft Defender and no virus was detected.

Dated: April 14, 2025

/s/ Mark S. Davies

Mark S. Davies

CERTIFICATE OF SERVICE

I hereby certify that I electronically filed the foregoing with the Clerk of the Court for the United States Court of Appeals for the Third Circuit by using the appellate CM/ECF system on April 14, 2025. All counsel of record are registered CM/ECF users, and service will be accomplished by the CM/ECF system.

Dated: April 14, 2025

/s/ Mark S. Davies

Mark S. Davies

APPENDIX

TABLE OF CONTENTS

	<u>Page</u>
Order Granting ROSS’s Motion for Certification of Interlocutory Appeal Pursuant to 28 U.S.C. § 1292(b), (D.E. 799)	App. 1
Memorandum Opinion re: Order Denying Plaintiffs’ Summary Judgment Motion for Copyright Infringement and Denying Defendant’s Motion for Fair Use, (D.E. 547).....	App. 2
Memorandum Opinion re: Order Granting Plaintiffs’ Renewed Summary Judgment Motion for Copyright Infringement and Denying Defendant’s Motion for Fair Use, (D.E. 770)	App. 36

ORDER: I GRANT Ross's motion for interlocutory appeal and stay pending appeal. D.I. [786]. Though I remain confident in my February 2025 summary judgment opinion, I recognize that there are substantial grounds for difference of opinion on controlling legal issues in this case. These issues have the potential to change the shape of the trial. I thus certify the following two questions to the Third Circuit: (1) whether the West headnotes and West Key Number System are original; and (2) whether Ross's use of the headnotes was fair use. A short opinion further explaining my reasoning will follow. Under the stay, the pretrial conference on April 9, 2025 and the trial the week of May 12, 2025 are postponed pending appeal. Ordered by Judge Stephanos Bibas on 04/03/2025. (jfm)

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

THOMSON REUTERS ENTERPRISE
CENTRE GMBH and WEST PUBLISH-
ING CORP.,

Plaintiffs,

v.

No. 1:20-cv-613-SB

ROSS INTELLIGENCE INC.,

Defendant.

Jack B. Blumenfeld, Michael J. Flynn, MORRIS, NICHOLS, ARSHT & TUNNELL LLP,
Wilmington, Delaware; Dale M. Cendali, Eric A. Loverro, Joshua L. Simmons,
KIRKLAND & ELLIS LLP, New York, New York.

Counsel for Plaintiffs

David E. Moore, Bindu A. Palapura, Andrew L. Brown, POTTER ANDERSON & COR-
ROON LLP, Wilmington, Delaware; Gabriel M. Ramsey, Warrington Parker, Joa-
chim B. Steinberg, Jacob Canter, Christopher J. Banks, Shira Liu, Margaux
Poueymirou, Anna Z. Saber, CROWELL & MORING LLP, San Francisco, California;
Mark A. Klapow, Lisa Kimmel, Crinesha B. Berry, CROWELL & MORING LLP, Wash-
ington, D.C.

Counsel for Defendant

MEMORANDUM OPINION

September 25, 2023

BIBAS, *Circuit Judge*, sitting by designation.

Facts can be messy even when parties wish they were not. But summary judgment is proper only if factual messes have been tidied. Courts cannot clean them up.

Thomson Reuters, a media company, owns a well-known legal research platform, Westlaw. It alleges that Ross, an artificial intelligence startup, illegally copied important content from Westlaw. Thomson Reuters thus seeks to recover from Ross. Both sides move for summary judgment on a variety of claims and defenses. But many of the critical facts in this case remain genuinely disputed. So I largely deny Thomson Reuters’s and Ross’s motions for summary judgment.

I. BACKGROUND

Many facts are disputed, but the basic story is not. Thomson Reuters’s Westlaw platform compiles judicial opinions according to its Key Number System. That system organizes opinions by the type of law. Westlaw also adds “headnotes”: short summaries of points of law that appear in the opinion. Each headnote is tied to a key number. Clicking on the headnote takes the user to the corresponding passage in the opinion. Clicking on the key number takes the user to a list of cases that make the same legal point. Westlaw has a registered copyright on its “original and revised text and compilation of legal material,” which includes its headnotes and Key Number System. D.I. 255-7, at 8.

Ross Intelligence is a legal-research industry upstart. It sought to create a “natural language search engine” using machine learning and artificial intelligence. D.I. 310, at 4. It wanted to “avoid human intermediated materials.” *Id.* Users would enter

questions and its search engine would spit out quotations from judicial opinions—no commentary necessary.

To leverage machine learning, Ross needed legal material to train the machine. At first, it tried to get a license to use Westlaw, but Thomson Reuters does not let users use Westlaw to develop a competing platform. So Ross turned to a third-party legal-research company, LegalEase Solutions. (LegalEase, in turn, hired a subcontractor, Morae Global. But the parties do not distinguish between LegalEase’s and Morae’s conduct, so I will refer only to LegalEase.)

Ross told LegalEase to create memos with legal questions and answers. The questions were meant to be those “that a lawyer would ask,” and the answers were direct quotations from legal opinions. D.I. 310, at 4. The so-called Bulk Memo Project produced about 25,000 question-and-answer sets. Each memo had one question plus four to six answers and rated each answer’s relevance. LegalEase created the memos both manually and, for a time, with the help of a text-scraping bot.

Ross says it converted the LegalEase memos into usable machine-learning training data. That involved first encoding the written language as numerical data and then running the data through a “Featurizer” that “performed various mathematical ... calculations on the text.” D.I. 272, at 8.

The core of this suit stems from the Bulk Memo Project. Thomson Reuters says the questions were essentially headnotes with question marks at the end. Ross admits that the headnotes “influence[d]” the questions but says lawyers ultimately drafted them, instead of copying them. D.I. 272, at 4–5. Though Thomson Reuters

contends that all 25,000 are copies, it has moved for summary judgment on just 2,830. It says LegalEase’s copying of those 2,830 is undisputed because Ross’s own expert admitted it.

Beyond the Bulk Memo Project, LegalEase provided Ross with two other relevant services. First, LegalEase sent Ross a list of 91 legal topics from Westlaw’s Key Number System. Ross admits that it “considered” these topics when creating its own set of 38 topics that were used in an experimental “Classifier Project.” D.I. 272, at 9–10. But it ultimately abandoned the Project. LegalEase also sent Ross 500 judicial opinions, including Westlaw’s headnotes, key numbers, and other annotations. Ross says it did nothing with these opinions.

In this opinion, I address five summary-judgment motions. Thomson Reuters has moved for summary judgment on its copyright-infringement claim (limited to the 2,830 memos mentioned), and both sides have moved for summary judgment on Ross’s fair-use defense. Thomson Reuters has also moved for summary judgment on its tortious-interference-with-contract claim, and Ross has counter-moved on its preemption defense to that claim.

“The court shall grant summary judgment if the movant shows that there is no genuine dispute as to any material fact and the movant is entitled to judgment as a matter of law.” Fed. R. Civ. P. 56(a). A dispute is “genuine” if a reasonable jury could resolve it in favor of either side. *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 250 (1986). And a fact is “material” if it “could affect the outcome.” *Lamont v. New Jersey*,

637 F.3d 177, 181 (3d Cir. 2011). I view the facts in the light most favorable to the nonmovant. *Id.* at 179 n.1.

II. COPYRIGHT INFRINGEMENT

A copyright-infringement claim has three elements: ownership of a valid copyright, actual copying, and substantial similarity. *See Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 361 (1991); *Dam Things from Denmark v. Russ Berrie & Co.*, 290 F.3d 548, 561–62 (3d Cir. 2002). Here, all three elements are at least partly disputed. But the dispute over the second element is legal, so I can decide it now. And because Ross hired LegalEase to do the copying (if there was any), Thomson Reuters also couches its argument in terms of direct, contributory, and vicarious liability. So after addressing the three infringement elements, I will consider each of these liability theories as well.

A. The parties still dispute breadth and validity of Westlaw's copyright

Ross bets a good chunk of its infringement defense on Westlaw's being registered as a compilation. Ross's theory is this: because Westlaw has just one copyright registration, comprising hundreds of thousands of headnotes and key numbers, copying a mere few thousand is not enough for infringement.

Ross's gamble does not pay off. A copyright in a compilation extends to the copyrightable pieces of that compilation. *Educ. Testing Servs. v. Katzman*, 793 F.2d 533, 538–39 (3d Cir. 1986) (abrogated on other grounds) (“The fact that a registrant designates the material as a compilation does not in itself signify that the constituent material is not also covered by the copyright.”). And when the author of a compilation presents facts through his own original words, “[o]thers may copy the underlying facts

from the publication, but not the precise words used to present them.” *Feist*, 499 U.S. at 348. Plus, though a plaintiff must have a registration to bring a federal suit for infringement, it can sue on all protected components of that one registration. 2 David Nimmer, *Nimmer on Copyright* § 7.16(B)(5)(c) (2023).

The cases Ross cites are the exceptions that prove the rule. In those cases, the copyright holder owned *only* the compilation. In one case cited, the plaintiff had a compilation copyright in the organization and selection of state legal forms. *Ross, Brouins & Oehmke, P.C. v. Lexis Nexis Grp.*, 463 F.3d 478, 480 (6th Cir. 2006). Though the underlying entries were in the public domain, the organizer’s exact selection and arrangement were copyrightable. *Id.* Even though the defendant copied and compiled 61% of the forms from the plaintiff’s compilation, there was no infringement because the defendant’s compilation was not the “same selection.” *Id.* at 483. So in these cases, the plaintiffs owned “thin” copyrights: other than their selection and arrangement choices, none of their compilations’ components were protectable.

Here, only the Key Number System aligns with the compilation caselaw: It is Westlaw’s method of organizing and arranging judicial opinions. So Thomson Reuters could have a valid copyright in this method of arrangement but not in the underlying opinions. That said, to qualify for copyright protection, “the manner of rearranging” and organizing the unprotectable underlying works “must constitute more than a minimal contribution.” Nimmer, *supra*, § 3.03(A). This “threshold for originality is low,” but the parties dispute facts needed to figure out if the System clears the bar. *Id.* § 3.04(B)(2)(a).

Thomson Reuters alleges that employees make creative organizing decisions to update and maintain the System and that the System is unique among its competition. But Ross replies that the System is unoriginal because 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.” D.I. 310, at 3. And although Thomson Reuters’s registered copyright could protect its Key Number System, the jury needs to decide its originality, whether it is in fact protected, and how far that protection extends.

In contrast, the headnotes are not aptly described by the compilation caselaw. Headnotes are just short written works, authored by Thomson Reuters, so they could receive standalone, individual copyright protection. *See* 17 U.S.C. § 103. This distinguishes Thomson Reuters’s copyright in its headnotes from the “thin,” compilation-only copyrights in Ross’s examples. So I must consider the alleged headnote copyright infringement at the level of each individual headnote, rather than at the level of the entire Westlaw compilation.

That said, Thomson Reuters’s allegedly original expression in its headnotes still reflects uncopyrightable judicial opinions. So the strength of its copyright depends on how much the headnotes overlap with the opinions. Closely hewing differs from copying: If a headnote merely copies a judicial opinion, it is uncopyrightable. But if it varies more than “trivial[ly],” then Westlaw owns a valid copyright. *L. Batlin & Son, Inc. v. Snyder*, 536 F.2d 486, 490 (2d Cir. 1976).

The parties dispute how Thomson Reuters develops its headnotes and how closely those headnotes resemble uncopyrightable opinions. Thomson Reuters points to evidence that its headnotes are original representations of its attorney-editors' views—summarizing the most important case facts, highlighting key issues, and describing the holdings. Ross, though, presents evidence that Thomson Reuters's protocols required headnotes to “follow or closely mirror the language of judicial opinions.” D.I. 272, at 13. This leaves a genuine factual dispute about how original the headnotes are. And this fact will serve double duty: it affects the strength and extent of Thomson Reuters's copyright, and it also goes to whether Ross was copying the headnotes or the opinions themselves.

In sum, I cannot decide the first element of Thomson Reuters's copyright infringement claim at summary judgment.

B. As a matter of law, Ross actually copied at least portions of the Bulk Memos

Next, Thomson Reuters must show that Ross (or LegalEase) “actually copied” its copyrighted work. “Actual copying focuses on whether the defendant did, in fact, use the copyrighted work in creating his own.” *Tanksley v. Daniels*, 902 F.3d 165, 173 (3d Cir. 2018). If Ross “truly created [its] work independently, then no infringement has occurred, irrespective of similarity.” *Id.* There are two ways to show actual copying: Thomson Reuters can present direct evidence. Or it can present circumstantial evidence demonstrating that Ross or LegalEase had access to the copyrighted work and that their work contains similarities probative of copying. *See id.*

Thomson Reuters presents both. LegalEase admitted to copying at least portions of the headnotes directly. As for circumstantial evidence, Ross does not dispute that LegalEase had access to Westlaw, which included access to headnotes. Though the similarities between Thomson Reuters's and Ross's work might not be substantial (that is a jury question), no reasonable jury could say that the similarities are not at least probative of some copying. And while Ross argues that any copying that occurred was miniscule in the grand scheme of the compilation, that framing misses the mark for the reasons given above. So Thomson Reuters has satisfied the actual-copying element as a matter of law.

C. Substantial similarity must go to the jury

The last element of direct infringement is substantial similarity. Substantial similarity asks whether “the ordinary observer, unless he set out to detect the disparities [in the two works], would be disposed to overlook them, and regard their aesthetic appeal as the same.” *Id.* at 174 (alteration in original) (quoting *Peter Pan Fabrics, Inc. v. Martin Weiner Corp.*, 274 F.2d 487, 489 (2d Cir. 1960) (Hand, J.)). In other words, I ask whether an ordinary person would view the two works as basically the same.

This case features several wrinkles in the substantial-similarity analysis. First, the Bulk Memos could appear similar to Thomson Reuters's headnotes because they share an underlying source: uncopyrightable judicial opinions. But I must determine whether Ross's work is substantially similar to Thomson Reuters's *protected* expression, not just the opinions. Second, we contextualize the ordinary-observer test. *See id.* at 172 n.3. And here, the ordinary consumers of both parties' products are lawyers.

So I should be attuned to differences a lawyer might notice that a layperson might not. Finally, the Third Circuit “has [generally] rejected the usefulness of experts in answering” the substantial-similarity question. *Id.* at 172. I thus do not give much weight to the parties’ dueling expert reports on this issue.

Substantial similarity “is usually an extremely close question of fact, which is why ... summary judgment has traditionally been disfavored in copyright litigation.” *Id.* at 171 (internal quotation marks omitted). Thomson Reuters argues that it can overcome this presumption because Ross’s expert allegedly made an “admission.” But this so-called admission does not get Thomson Reuters over the summary-judgment line.

The “admission” goes as follows: Ross’s expert compiled the 25,000-plus Bulk Memo questions. She then paired each question with the Westlaw headnote most like it and paired each headnote with the judicial opinion passage most like the headnote. Next, on a scale of one to five, she rated two things: the similarity between the question and the headnote and the similarity between the headnote and the judicial opinion passage. The “admission” Thomson Reuters refers to is the 2,830 instances in which a question was rated as a close match to a post-1927 headnote, but that headnote was rated as not a verbatim or near-verbatim copy of judicial opinion text. (Copyrighted works created before 1927 are in the public domain and are not protected.)

Ross disagrees, breaking the headnotes into three groups. First, it says Thomson Reuters did not identify 1,623 of these 2,830 headnotes in its supplemental response, so they are not part of the case. Second, and more substantively, Ross says 1,019 questions are nearly identical to judicial opinions. Finally, it says nothing about

the remaining 188 entries, other than that they make up a tiny fraction of Westlaw's compilation. I take each group in turn.

First, I agree that Thomson Reuters did not identify the 1,623 headnotes. In an earlier Order, I reminded Thomson Reuters of its burden to show infringement and told it to identify "what, precisely, was copied." D.I. 201, at 1. Its response identified thousands of allegedly copied headnotes. But these 1,623 headnotes were not among those specifically identified. Thomson Reuters argues that it identified the Bulk Memos in which these headnotes were copied and that they incorporate by reference the cases in which the headnotes appear. But that is not precise enough. Producing the cases with the allegedly copied headnotes was what prompted Ross's objection and my Order to be more specific in the first place. So Thomson Reuters is limited (for purposes of summary judgment) to the 1,207 headnotes specifically identified.

There are genuine factual disputes over the second group. In her report, the Ross expert said each of these 1,019 questions had high overlap with a headnote and that the headnote was not identical to opinion text. But she did not—and could not—take a position on whether the headnotes and questions were "substantially similar" under the ordinary-observer test. And more specifically, the report does not pinpoint how much similarity came solely from Thomson Reuters's protected expression.

Plus, Ross offers contrary evidence for these 1,019 entries. It shows that either the judicial opinion text is identical to the headnote or that the opinion text is more similar to the Bulk Memo question than the headnote is to the question. This supports the contention that similarity between Ross's and Thomson Reuters's work

stems from uncopyrightable judicial opinions, rather than from Thomson Reuters's original expression.

Thomson Reuters objects that Ross did not disclose its expert's methodology. But substantial similarity is not especially scientific: the question boils down to "good eyes and common sense." *Petrella v. Metro-Goldwyn-Mayer, Inc.*, 572 U.S. 663, 684 (2014). This makes Ross's mode of argument valid. Because both sides have lobbed conflicting expert reports (which I give little weight anyways) at each other and because a reasonable jury could agree with either side, I must send the question of substantial similarity for these 1,019 entries to the jury, where it typically belongs.

Finally, Ross does not object to the remaining 188 entries. Though substantial similarity is usually a close factual question, I will not review each of the 188 entries and make arguments for Ross. And for the reasons above, its reliance on Westlaw's copyright being solely in a compilation is misplaced. So each of these headnotes is substantially similar to its associated Bulk Memo question. But as noted above, whether this copying constitutes infringement depends on whether these headnotes are protected expression. And that rests on factual determinations the jury still must make. Plus, to recover from Ross, Thomson Reuters must win on one of its theories of liability and defeat Ross's fair-use defense.

D. All of Thomson Reuters's theories of infringement liability must go to trial

1. *Direct liability.* Thomson Reuters's theory of Ross's direct liability is uncontested: Ross hosted copies of the Bulk Memos on its servers, copied the content into its machine-learning "portal," transmitted another copy to a different server, created

more copies on employees' computers, then processed and labeled them by copying parts into another document. D.I. 250, at 13. Simply hosting a copy on a server might not seem like copying, but it is. *See MAI Sys. Corp. v. Peak Comput. Inc.*, 991 F.2d 511 (9th Cir. 1993); Nimmer, *supra*, § 8.08(A).

The unstated premise of this theory is that Ross violated Westlaw's reproduction right by making copies of the Bulk Memos. So for Thomson Reuters to succeed on direct liability, LegalEase's Bulk Memos must be unauthorized copies of protected expression. For making a copy of a non-copy is not copyright infringement. But because whether the Bulk Memos copied protected expression depends on factual determinations the jury must make, I cannot resolve direct liability at summary judgment.

2. *Contributory liability.* For Ross to be contributorily liable, Thomson Reuters must show that Ross (1) knew LegalEase was infringing and (2) materially contributed to or induced that infringement. *See Leonard v. Stemtech Int'l Inc.*, 834 F.3d 376, 387 (3d Cir. 2016). The parties dispute both prongs.

At best, Thomson Reuters has strong evidence that Ross knew LegalEase was using Westlaw. But knowledge or even encouragement to use Westlaw is not enough. One might expect a legal-research project to be completed using Westlaw, but merely using the service is not infringement. Plus, Ross points to evidence that it did not know LegalEase was infringing and never specifically instructed LegalEase to use Westlaw. Thomson Reuters has not done enough to prove that Ross knew about and

materially contributed to LegalEase's infringement. So this is not proper for summary judgment.

Thomson Reuters tries to bridge the gap between use and infringement by arguing that LegalEase breached its Westlaw license and Ross knew it. Once LegalEase breached the license, Thomson Reuters says, everything it did on Westlaw was copyright infringement. So Ross's knowledge of LegalEase's breach confers on Ross knowledge of the infringement. This argument mangles the interaction between licenses and copyright infringement.

In many copyright cases, licenses are used as a defense. In cases involving a license defense, one party claims infringement, and the other side claims they had permission through the license. But if the side claiming permission exceeded the scope of the license, it can be liable for infringement. *MacLean Assocs, Inc. v. Wm. M. Mercer-Meidinger-Hansen, Inc.*, 952 F.2d 769, 779 (3d Cir. 1991). That said, the copyright owner's rights must still be infringed by the specific activity that exceeds the scope of the license. Otherwise, the owner must litigate the license violations as breach-of-contract claims. *See S.O.S., Inc. v. Payday, Inc.*, 886 F.2d 1081, 1088–90 (9th Cir. 1989).

Here, there is no real dispute that LegalEase and Ross's alleged copying was not protected by a license. Rather, the issue is whether their actions constitute infringement of Thomson Reuters's copyright protections. And the license issue is irrelevant to proving Ross's contributory liability for LegalEase's infringement. So I deny summary judgment on the contributory-liability theory.

3. *Vicarious liability.* For vicarious liability, Thomson Reuters must show that Ross had “(1) the right and ability to supervise or control the infringing activity; and (2) a direct financial interest in such activities.” *Leonard*, 834 F.3d at 388. Taking the elements in reverse, Ross does not contest that it had a financial interest in the alleged copies—it used the Bulk Memos to train AI, its core product. But it does contest whether it could supervise LegalEase. The control element is a matter of “practical ability.” *Id.* at 389. So the determination is often fact-intensive. Evidence needs to support a finding that the defendant was “in a position to police the direct infringer[.]” *Id.* at 388 (quoting *Fonovisa, Inc. v. Cherry Auction, Inc.*, 76 F.3d 259, 262–63 (9th Cir. 1996)). Thomson Reuters has testimony saying that Ross dictated LegalEase’s practices. But Ross pushes back with evidence that LegalEase was secretive and resisted micromanagement. Thus, whether Ross had the “practical ability” to control LegalEase’s “infringing activity” remains a disputed factual question for the jury to resolve.

III. FAIR USE MUST GO TO A JURY

The parties have cross-moved on Ross’s fair-use defense. Fair use balances four factors: (1) the purpose and character of the use, (2) the nature of the copyrighted work, (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole, and (4) the effect of the use upon the potential market for the copyrighted work. 17 U.S.C. §107. The first and fourth factors are most important. *See Authors Guild v. Google, Inc.*, 804 F.3d 202, 213–14 (3d Cir. 2015).

Fair use is a mixed question of law and fact. Though applying the test “primarily involves legal work,” it requires “determination of subsidiary factual questions” about

the copying or the marketplace. *Google LLC v. Oracle Am., Inc.*, 141 S. Ct. 1183, 1199–200 (2021). Here, all of this must go to a jury.

A. The purpose and character of the use will be determined by contested facts

This first factor has two subparts: commerciality and transformativeness. *See Authors Guild*, 804 F.3d at 214, 218–19. (Bad faith is a minor subpart, also typically filed under this factor, and I will address it at the end.) Commercial use weighs against finding fair use, while transformative use weighs in favor. *Id.* at 218–20. And these considerations interact. “The more transformative the new work, the less will be the significance of ... commercialism....” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994). Commerciality is straightforward: it asks whether the use was for profit. *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 562 (1985). Transformativeness is less so: “a transformative use is one that communicates something new and different from the original or expands its utility, thus serving copyright’s overall objective of contributing to public knowledge.” *Authors Guild*, 804 F.3d at 214.

Ross’s uses were undoubtedly commercial. And one of its goals was to compete with Westlaw. Thomson Reuters contends that this commercial use weighs heavily against finding fair use. In support of this, it cites the Supreme Court’s recent decision in *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 143 S. Ct. 1258 (2023). There, the Court determined that the use in question was not fair largely by emphasizing its commercial nature. *See id.* at 1279–80. But I decline to overread one decision, especially because the Court recognized that “use’s transformativeness

may outweigh its commercial character” and that in *Warhol*, “both elements point[ed] in the same direction.” *Id.* at 1280. Plus, just two terms ago, in a technological context much more like this one, the Court placed much more weight on transformation than commercialism. *Google*, 141 S. Ct. at 1204 (“[A] finding that copying was not commercial in nature tips the scales in favor of fair use. But the inverse is not necessarily true, as many common fair uses are indisputably commercial.”). So I focus on transformativeness.

Thomson Reuters paints a black-and-white picture on transformativeness: Westlaw is a legal-research platform that synthesizes the law; Ross used Westlaw’s syntheses to build a legal-research platform that also synthesizes the law. Ross, on the other hand, presents a more nuanced account: Westlaw headnotes and key numbers annotate opinions for users. Ross wanted to build a search engine that “avoids human intermediated materials,” meaning a user would simply enter a query and get a responsive quotation from a judicial opinion, no clicking around or commentary needed. D.I. 272, at 1–3. Though Ross and Westlaw both help answer legal questions, Ross says it transformed the Westlaw headnotes beyond recognition.

Ross describes its process of transforming the Bulk Memos like this: First, it receives the Bulk Memos in its database. Then, it converts the plain-language entries into numerical data. Next, it feeds that data into its machine-learning algorithm to teach the artificial intelligence about legal language. The idea is that the artificial intelligence will be able to recognize patterns in the question-answer pairs. It can

then use those patterns to find answers not just to the exact questions fed into it, but to all sorts of legal questions users might ask.

Ross says that the caselaw on “intermediate copying” most appropriately reflects its use. In those cases, the users copied material to discover unprotectable information or as a minor step towards developing an entirely new product. So the final output—despite using copied material as an input—was transformative. In *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), the defendant copied Sega’s copyrighted software. But it did so only to figure out the functional requirements to make games compatible with Sega’s gaming console. *Id.* at 1522. That functional information was unprotected, so the copying was fair use. *Id.* at 1522–23.

Similarly, in *Sony Computer Entertainment Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000), the defendant used a copy of Sony’s software to reverse engineer it and create a new gaming platform on which users could play games designed for Sony’s gaming system. *Id.* at 601. The court concluded that this was fair use for two reasons: the defendant created “a wholly new product, notwithstanding the similarity of uses and functions” between it and Sony’s system, and the “final product [did] not itself contain infringing material.” *Id.* at 606. The Supreme Court has cited these intermediate copying cases favorably, particularly in the context of “adapt[ing] the doctrine of fair use ... in light of rapid technological change.” *Google*, 141 S. Ct. at 1198 (cleaned up).

Thomson Reuters says the intermediate-copying cases are inapt. It argues that whereas in those cases, the copiers sought to “study functionality or create compatibility,” here Ross simply sought to “train[] its AI” by “cop[ying] the creative decisions of West[law]’s attorney-editors precisely because it wanted to replicate them.” D.I. 317, at 11. And it contends that Ross merely translated the headnotes into numerical data and that translation is “paradigmatic derivative work[.]” D.I. 317, at 10.

But Ross says its AI studied the headnotes and opinion quotes only to analyze language patterns, not to replicate Westlaw’s expression. So the translation was only a minor step in a broader, transformative use. *See Sega*, 977 F.2d at 1514–15, 1518–19 (holding that, though programmers wrote down and translated Sega’s object code, these acts were a minor step towards a transformative use). If Ross’s characterization of its activities is accurate, it translated human language into something understandable by a computer as a step in the process of trying to develop a “wholly new,” albeit competing, product—a search tool that would produce highly relevant quotations from judicial opinions in response to natural language questions. This also means that Ross’s final product would not contain or output infringing material. Under *Sega* and *Sony*, this is transformative intermediate copying.

So whether the intermediate copying caselaw tells us that Ross’s use was transformative depends on the precise nature of Ross’s actions. It was transformative intermediate copying if Ross’s AI only studied the language patterns in the headnotes to learn how to produce judicial opinion quotes. But if Thomson Reuters is right that Ross used the untransformed text of headnotes to get its AI to replicate and reproduce

the creative drafting done by Westlaw’s attorney-editors, then Ross’s comparisons to cases like *Sega* and *Sony* are not apt. Again, this is a material question of fact that the jury needs to decide.

Finally, the parties clash over whether Ross’s use was in bad faith. But bad faith is at most a minor consideration in the fair use analysis. Indeed, the Supreme Court has expressed skepticism about whether it has any role to play at all. *Google*, 141 S. Ct. at 1204. And bad faith is particularly unimportant here. Thomson Reuters argues that Ross demonstrated bad faith by initially asking to license Westlaw, being denied, and then hiring LegalEase to illicitly gain access to it. But the Supreme Court has foreclosed this line of reasoning, explaining that “[i]f the use is otherwise fair, then no permission need be sought or granted. Thus, being denied permission to use a work does not weigh against a finding of fair use.” *Campbell*, 510 U.S. at 585 n.18. So I can ignore bad faith. And the first fair use factor comes down to the jury’s finding of formativeness.

B. The nature of the copyrighted work favors fair use, but factual questions remain

The second factor asks about the nature of the copyrighted work. The work gets more protection, and copies are less likely to be fair, if it is near the “core of intended copyright protection.” *Id.* at 586. But “[t]he scope of fair use is greater when ‘informational’ as opposed to more ‘creative’ works are involved.” *Hustler Mag. Inc. v. Moral Majority Inc.*, 796 F.2d 1148, 1153–54 (9th Cir. 1986). So although judges should not act as critics, we consider “whether the work was creative, imaginative, and original.” *MCA, Inc. v. Wilson*, 677 F.2d 180, 182 (2d Cir. 1981). That said, “[t]he second factor

has rarely played a significant role in the determination of a fair use dispute.” *Authors Guild*, 804 F.3d at 220.

The analysis for this factor mirrors much of my earlier discussion of the validity and strength of Thomson Reuters’s copyright. As explained above, this depends largely on factual questions that the jury must decide, so I cannot resolve this factor at summary judgment.

But I will note here that the Key Number System is far from the core of copyright. Even if the system involves making creative decisions about how to organize opinions and other material and is an original method of organization, it is merely a way to arrange “informational” material. So the system inherently involves significantly less creative or original expression than traditionally protected materials, such as literary works or visual art, and is much less “imaginative.”

The headnotes are closer, but still not especially close to the core. “The law generally recognizes a greater need to disseminate factual works than works of fiction or fantasy.” *Harper*, 471 U.S. at 563. And though editors may have made creative choices about which points of law to summarize, how to summarize them, and where to attach the headnote, those choices are constrained. In general, the headnotes will flag the most salient points of law, largely track the language of the opinion, and be placed at the beginning of a paragraph. This approach is akin to news reporting, which, though protected, must be carefully separated from the unprotected underlying facts. *See Author’s Guild*, 804 F.3d at 220. So, although a jury must decide how closely headnotes reflect the language of judicial opinions and, in turn, precisely how much

protection they are afforded, they are not at the core of intended copyright protection. Thus, although an ultimate decision on factor two must wait until trial, this factor seems to favor fair use.

C. The amount and substantiality of the copying depends on the nature of Ross’s AI outputs

Third, I consider the amount of copying as well as whether the copying took the original work’s “heart.” *Campbell*, 510 U.S. at 589.

Defining the work at issue matters in determining the amount of copying done. If we define it at the level of each headnote, the copying was allegedly completed for some 25,000 headnotes. If we define it at the level of the compilation, however, the copying was less substantial, though headnotes likely represent the “heart” of Westlaw’s expression.

And defining the use is again important because “even a small amount of copying may fall outside of the scope of fair use where the excerpt copied consists of the heart of the original work’s creative expression.” *Google*, 141 S. Ct. at 1205 (internal quotation marks omitted). Conversely, “copying a larger amount” can still be fair use “where the material copied captures little of the material’s creative expression.” *Id.* Plus, “[t]he ‘substantiality’ factor will generally weigh in favor of fair use where ... the amount of copying was tethered to a valid, and transformative, purpose.” *Id.* In particular, verbatim intermediate copying has consistently been upheld as fair use if the copy is “not reveal[ed] ... to the public.” *Authors Guild*, 804 F.3d at 221; *see also A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 638–640, 642 (4th Cir. 2009).

Here, the best definition is at the level of each headnote. As mentioned, the compilation registration also covers individually copyrightable materials. And each headnote counts. But the heart of each headnote is its original expression, not its link to the part of the opinion it summarizes. So if Ross’s AI works the way that it says, it is likely fair use because it produces only the opinion, not the original expression. “It cannot be said that a revelation is ‘substantial’ in the sense intended by the statute’s third factor if the revelation is in a form that communicates little of the sense of the original.” *Authors Guild*, 804 F.3d at 223.

Yet this factor also requires jury fact-finding. How Ross’s AI works and what output it produces remain disputed. The parties also fight over whether the use was “tethered to a valid ... purpose.” Westlaw says Ross copied far more than it needed. Ross says it needed a vast, diverse set of material to train its AI effectively. Though Ross need not prove that each headnote was strictly necessary, it must show that the scale of copying (if any) was practically necessary and furthered its transformative goals. So the third factor hinges on the answers to these disputed factual questions which the jury needs to resolve.

D. I cannot yet determine the effect of the use upon the market for the work

Finally, factor four asks whether the use had a “meaningful or significant effect” on the value of the original or its potential market. *Authors Guild*, 804 F.3d at 224. And “[t]his inquiry must take account not only of harm to the original but also of harm to the market for derivative works.” *Harper*, 471 U.S. at 568. Yet not all losses are created equal. I must also consider the “source of the loss.” *Google*, 141 S. Ct. at

1206. Again, we come back to the fundamental premise that copyright protects expression. If the source of the loss is not that the original’s expression is being appropriated, “the type of loss of sale envisioned above will generally occur in relation to interests that are not protected by the copyright.” *Authors Guild*, 804 F.3d at 224.

And transformativeness feeds into this factor as well. “[T]he more the copying is done to achieve a purpose that differs from the purpose of the original, the less likely it is that the copy will serve as a satisfactory substitute for the original.” *Id.* at 223 (citing *Campbell*, 510 U.S. at 591). Finally, in evaluating market impact, courts must pay special attention to “the realities of how technological works are created and disseminated.” *Google*, 141 S. Ct. at 1199.

Here, those “realities” are disputed. Thomson Reuters claims three potential markets, but they boil down to two: the market for Westlaw itself as a legal research platform and the market for its data. It says Ross’s plan all along was to create a substitute for Westlaw. And it says that this plan worked, as some Ross customers cancelled their Westlaw subscriptions. As for the market for its data, Thomson Reuters says there is a traditional licensing market and a burgeoning one for AI training data. It argues that it lost traditional licensing revenue because Ross obtained Westlaw content through LegalEase. And it suggests that there is a potential market for Westlaw’s training data; after all, Ross paid LegalEase over a million dollars for the Bulk Memos. That burgeoning market would be harmed by copying like Ross’s.

One fact is undisputed here: Ross and Thomson Reuters both compete in the market for legal research platforms. But that alone does not reveal whether Ross’s AI

product is a substitute for Westlaw. Ross’s use might be transformative, creating a brand-new research platform that serves a different purpose than Westlaw. If so, it is not a market substitute. Ross also argues that Thomson Reuters has never participated—and would never participate—in this market for its training data. Because a reasonable jury could find for either side on these factual market-impact questions, I cannot resolve them at summary judgment.

Finally, “we must take into account the public benefits the copying will likely produce.” *Google*, 141 S. Ct. at 1206. And “we are free to consider the public benefit resulting from a particular use notwithstanding the fact that the alleged infringer may gain commercially.” *Sega*, 977 F.2d at 1523. This “[p]ublic benefit need not be direct or tangible, but may arise because the challenged use serves a public interest.” *Id.*

The parties provide competing narratives of public benefit. Ross’s research platform might increase access to the law at a lower cost. Or it might just reduce the incentives for Thomson Reuters, and similarly situated entities, to create content like headnotes in the future.

Deciding whether the public’s interest is better served by protecting a creator or a copier is perilous, and an uncomfortable position for a court. Copyright tries to encourage creative expression by protecting both. Here, we run into a hotly debated question: Is it in the public benefit to allow AI to be trained with copyrighted material?

The value of any given AI is likely to be reflected in the traditional factors: How transformative is it? Can the public use it for free? Does it discourage other creators

Grace Consulting, 307 F.3d 197, 217–18 (3d Cir. 2002). As its name suggests, that test asks whether the state-law claim has an element that a §106 claim would not. *Id.* at 217.

The test is easy to say but hard to apply. In some sense, every claim other than a state copyright claim has some extra element. For example, although almost every common law claim requires damages, §106 does not because federal law supplies statutory damages. But that alone does not make the claims meaningfully different. So courts have modified the test, asking whether the extra element makes the claim qualitatively different. *Id.* To avoid preemption, the “gravam[e]n” of the state claim must differ from one of the §106 rights. *Id.* at 218 (quoting *Computer Assocs. Int’l, Inc. v. Altai, Inc.*, 982 F.2d 693, 717 (2d Cir. 1992)).

Ross says tortious-interference claims are almost always preempted. Though it cites good authority for that argument, it takes that authority out of context. A tortious-interference claim that says something like, “You copied our work, thus interfering with our contracts to license that work” is preempted because it merely identifies one of the consequences of a §106 violation. But some of Thomson Reuters’s claims are different. It brings some of its claims as tortious-interference claims—rather than as standard breach-of-contracts claims—solely because it seeks to hold Ross liable for the acts of a third party, LegalEase. So the preemption analysis should look more like it would with a typical breach-of-contract claim than with a claim that just identifies a consequence of copyright infringement.

To sum up, a claim is preempted if (1) the material is within the subject matter of copyright and (2) the gravamen of the claim is equivalent to a § 106 right.

The anti-competition claim is preempted. Thomson Reuters’s first tortious-interference claim concerns this provision:

You may not sell, sublicense, distribute, display, store or transfer [West’s] products or any data in [its] products in bulk or in any way that could be used to replace or substitute for [its] products in whole or in part or as a component of any material offered for sale, license or distribution to third parties.

D.I. 316, at 4 (alterations in original).

Thomson Reuters says Ross induced LegalEase to breach this provision by hiring it to create the Bulk Memos and other materials sent to Ross. Thomson Reuters does not contest that this covers material that is within the subject matter of copyright. But it says the rights implicated are different. Not so.

The gravamen of this claim is the same as that of Thomson Reuters’s copyright claim. And the contract provision itself secures equivalent—indeed, sometimes identical—rights: The rights to sell, sublicense, distribute, and transfer are covered by § 106(3). The right to display is covered by § 106(5). The “in bulk” and “in any way that could be used to replace or substitute” phrases are incorporated in the fair-use analysis. And the “as a component of” language also tracks fair use and the derivative-work right under § 106(2).

Though this contract provision is framed in terms of competition, it is focused on one potential competitive threat: copying. That concern is the domain of federal law. So this first claim is preempted by the Copyright Act.

Thomson Reuters tries to analogize its provision to the ones at issue in cases like *Altera Corp. v. Clear Logic, Inc.*, 424 F.3d 1079 (9th Cir. 2005), and *Wellness Publishing v. Barefoot*, No. 02–3773, 2008 WL 108889 (D.N.J. Jan. 9, 2008). But those cases involved restrictions on use. Restricting a user’s use of copyrighted material is different from limiting the user’s ability to copy it. The latter is covered, and thus preempted, by the Copyright Act.

The anti-bot and password sharing claims are not preempted. The two other tortious-interference claims involve Westlaw’s anti-bot and password-sharing provisions:

You may not run or install any computer software or hardware on [West’s] products or network or introduce any spyware, malware, viruses, Trojan horses, backdoors or other software exploits.

Your access to certain products is password protected. You are responsible for assigning the passwords and maintaining password security. Sharing passwords is strictly prohibited.

D.I. 316, at 4–5 (alteration in original).

These provisions are not equivalent to § 106 rights. Unlike the competition provision, they govern use and manipulation of the site. Using a bot to scrape content might copy material in bulk. And a claim based on the harm from that copying itself would be preempted. But a claim based on simply introducing malware, independent of that malware’s goals, is not equivalent to any right in § 106. Likewise, a site might ban password sharing because they want to limit copying risk. But putting limits on access to the site is a separate restriction. Whether the material behind the password protection is copyrighted or not, the creator can protect the material for which it

charges users. Section 106 has nothing to say about that limit. So Thomson Reuters's second and third tortious-interference claims survive preemption.

B. Thomson Reuters's two remaining tortious-interference claims are partially disputed

I now consider the two surviving tortious-interference claims on the merits. Thomson Reuters must prove five elements: (1) there was a contract between LegalEase and Westlaw, (2) Ross knew about the contract and its terms, (3) Ross's intentional act was a significant factor causing the breach, (4) Ross had no justification, and (5) the breach harmed Thomson Reuters. *See WaveDivision Holdings, LLC v. Highland Cap. Mgmt., L.P.*, 49 A.3d 1168, 1174 (Del. 2012). (Earlier in this case, the parties disputed whether Minnesota or Delaware law applied. But I need not address this choice-of-law question because Minnesota's tortious-interference elements are the same as Delaware's. So I address each element below under Delaware law.)

1. There was a contract. For both claims, there is no dispute that the first element is met—there was a contract between Westlaw and LegalEase. But three of the four remaining elements involve genuine factual disputes that remain for trial.

2. It is unclear how much Ross knew. To satisfy the second element, Ross must have had actual or imputed knowledge of the substance of the contract rights, even if it did not know about the exact terms. *WaveDivision*, 49 A.3d at 1176. There is substantial record evidence that Ross knew at least something about Westlaw's contracting practices: Ross itself tried to enter a contract with Westlaw. An investor sent Ross a copy of the Westlaw terms and conditions. A Ross executive posed as an individual

practitioner to see the terms of use. And there are email exchanges in which Ross executives discuss specific provisions of the contract.

But Ross's evidence introduces some ambiguity. It disputes the timeline, saying that much of Thomson Reuters's evidence comes from well after Ross dealt with LegalEase. And it says that although it saw the contracts Westlaw offered it, it never saw Westlaw's contract with LegalEase. Westlaw's agreements can be tailored, Ross says, and it only ever saw the Canadian, not United States, agreement. Thomson Reuters counters that the agreements are materially the same, publicly disclosed, and seldom if ever altered. Whether this evidence, taken together, rises to the level of knowledge of the substance of the anti-bot and password-sharing provisions is a jury question.

3. *Ross may have intentionally caused a breach.* Ross met the third element if it (1) intended to interfere or (2) intended to reach a result with knowledge that it would interfere with the contract, even if interference was not the main purpose. Restatement (Second) of Torts § 766 cmt. j (Am. L. Inst. 1965). Ross intended for LegalEase to produce the Bulk Memos. And it was likely aware that LegalEase was using Westlaw. But whether Ross knew LegalEase was breaching or going to breach by using a bot or sharing passwords is far less clear. As explained in the discussion of indirect liability, each side has evidence suggesting different levels of Ross's involvement, control, and knowledge. So this element too is unfit for summary judgment.

4. *Whether Ross acted without justification is a factual question.* Courts commonly refer to the fourth element as doing something "not ... sanctioned by the rules of the

game.” *Avaya Inc., RP v. Telecom Labs, Inc.*, 838 F.3d 354, 383 (3d Cir 2016) (internal quotations omitted). Thomson Reuters said Ross did that. According to Thomson Reuters, Ross first sought a license from Westlaw, but when that rule-abiding approach failed, it hired LegalEase to end-run Thomson Reuters’s terms. Yet end-running or exploiting a loophole is not necessarily the same as violating the rules of the game. And whether Ross acted without justification largely depends on how it acted—which is a matter of dispute under the second and third elements. Plus, this element is typically a question of fact: Delaware law considers no less than seven factors, nearly all of which involve some underlying question of fact. *See WaveDivision*, 49 A.3d at 1174. So this element remains for trial.

5. *Thomson Reuters has shown harm.* The fifth element requires harm. Thomson Reuters says it lost subscription fees when LegalEase used a bot and shared passwords. If LegalEase did not have the help of a bot or multiple employees sharing one account, it would have had to buy more subscriptions or keep its subscriptions open for longer periods. This harm is distinct from the harm from copying: assuming copying was going to happen, Westlaw at least wanted to get paid while LegalEase did it. The bot and password sharing made the copying more efficient and cheaper, depriving Westlaw of fees. Ross does not contest this element, other than arguing that Thomson Reuters’s damages here are the same as the damages it would get for its copyright infringement claim. That misses the mark, so there is no genuine dispute.

In sum, Thomson Reuters is entitled to partial summary judgment on the first and fifth elements of tortious interference by using a bot and sharing passwords, but elements two through four remain for trial.

V. OTHER DEFENSES FAIL

Ross throws several other defenses at the wall, but none sticks. First, it no longer presses its First Amendment or first-sale defenses. Second, it raises laches. But laches does not apply to the copyright claim. *See Petrella*, 572 U.S. at 679. Plus, Thomson Reuters brought its tortious-interference claim promptly. Ross says that Thomson Reuters sat on its right to sue for years, but this is unsupported, and Ross cites no law saying that was too long.

Third, for its defenses of consent, waiver, estoppel, acquiescence, and license, Ross points to a fair-use provision in Westlaw’s terms of use. That provision just begs the fair-use question but does not provide an independent defense. Fourth, Ross alleges tort of another, saying it was not a “substantial factor” in the harm that LegalEase allegedly caused. D.I. 318, at 20. But again that argument is the same as its argument against the elements of the tortious-interference claim. And it has presented no evidence specifically for this defense, so I will not repackage other evidence for it. Instead, Ross can focus on defeating the tortious-interference claim’s elements directly.

Finally, Ross alleges lack of ownership because the headnotes are identical to public law. But Thomson Reuters has provided its registrations, and the extent of its expression is fully explored under the infringement and fair-use claims. Indeed, whether “lack of ownership” is even an affirmative defense is dubious—it seems to go to the first element of an infringement claim (ownership of a valid copyright). So I

grant summary judgment to Thomson Reuters on these miscellaneous affirmative defenses.

* * * * *

Thomson Reuters alleges that Ross copied protected aspects of Westlaw, both directly and indirectly through LegalEase. And Ross disputes almost all of Thomson Reuters's story. But it is not my role at summary judgment to sort through the evidence and tidy these factual messes. It is the jury's role at trial. So, with the small exceptions noted throughout this opinion, I deny both Ross's and Thomson Reuters's motions for summary judgment.

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

THOMSON REUTERS ENTERPRISE
CENTRE GMBH and WEST PUBLISH-
ING CORP.,

Plaintiffs,

v.

No. 1:20-cv-613-SB

ROSS INTELLIGENCE INC.,

Defendant.

Jack B. Blumenfeld, Michael J. Flynn, MORRIS, NICHOLS, ARSHT & TUNNELL LLP, Wilmington, Delaware; Dale M. Cendali, Eric A. Loverro, Joshua L. Simmons, KIRKLAND & ELLIS LLP, New York, New York; Yungmoon Chang, KIRKLAND & ELLIS LLP, Los Angeles, California; Miranda D. Means, KIRKLAND & ELLIS LLP, Boston, Massachusetts.

Counsel for Plaintiffs.

David Ellis Moore, Bindu Ann George Palapura, POTTER ANDERSON & CORROON LLP, Wilmington, Delaware; Jordan Ludwig, Emily T. Kuwahara, CROWELL & MORING LLP, Los Angeles, California; Ryan Henry Seewald, CROWELL & MORING LLP, Denver Colorado; Warrington Parker, Joachim B. Steinberg, Jacob Canter, Christopher J. Banks, Anna Z. Saber, Margaux Poueymirou, CROWELL & MORING LLP, San Francisco, California; Keith J. Harrison, Mark A. Klapow, Lisa Kimmel, Crinesha B. Berry, CROWELL & MORING LLP, Washington, D.C.

Counsel for Defendant.

MEMORANDUM OPINION

February 11, 2025

BIBAS, *Circuit Judge*, sitting by designation.

A smart man knows when he is right; a wise man knows when he is wrong. Wisdom does not always find me, so I try to embrace it when it does—even if it comes late, as it did here.

I thus revise my 2023 summary judgment opinion and order in this case. *See* Fed. R. Civ. P. 54(b); D.I. 547, 548; *Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc.*, 694 F. Supp. 3d 467 (D. Del. 2023). Now I (1) grant most of Thomson Reuters’s motion for partial summary judgment on direct copyright infringement and related defenses, D.I. 674; (2) grant Thomson Reuters’s motion for partial summary judgment on fair use, D.I. 672; (3) deny Ross’s motion for summary judgment on fair use, D.I. 676; and (4) deny Ross’s motion for summary judgment on Thomson Reuters’s copyright claims, D.I. 683.

I. ROSS MAKES A LEGAL AI TOOL AND WESTLAW’S OWNER SUES

The law is no longer a brooding omnipresence in the sky; it now dwells in legal-research platforms. Thomson Reuters owns one of the biggest of those platforms: Westlaw. D.I. 752-1 at 4. Users can pay to access its contents, including “case law, state and federal statutes, state and federal regulations, law journals, and treatises.” *Id.* “Westlaw also contains editorial content and annotations,” like the headnotes here. *Id.* Those headnotes summarize key points of law and case holdings. Westlaw organizes its content using the Key Number System, a numerical taxonomy. *Id.* Thomson Reuters owns copyrights in Westlaw’s copyrightable material. *Id.*

Ross, a new competitor to Westlaw, made a legal-research search engine that uses artificial intelligence. *Id.* To train its AI search tool, Ross needed a database of legal questions and answers. *Id.* at 5. So Ross asked to license Westlaw’s content. *Id.* But because Ross was its competitor, Thomson Reuters refused. *Id.* at 4–5.

So to train its AI, Ross made a deal with LegalEase to get training data in the form of “Bulk Memos.” *Id.* at 5. Bulk Memos are lawyers’ compilations of legal questions with good and bad answers. LegalEase gave those lawyers a guide explaining how to create those questions using Westlaw headnotes, while clarifying that the lawyers should not just copy and paste headnotes directly into the questions. D.I. 678-36 at 5–9. LegalEase sold Ross roughly 25,000 Bulk Memos, which Ross used to train its AI search tool. *See* D.I. 752-1 at 5; D.I. 769 at 30 (10:48:35). In other words, Ross built its competing product using Bulk Memos, which in turn were built from Westlaw headnotes. When Thomson Reuters found out, it sued Ross for copyright infringement.

In 2023, I largely denied Thomson Reuters’s motions for summary judgment on copyright infringement and the fair-use defense, and the case moved ahead toward trial. D.I. 547, 548. In the run-up to the August 2024 trial date, I studied the case materials more closely and realized that my prior summary-judgment ruling had not gone far enough. So I continued the trial and invited the parties to renew their summary-judgment briefing. D.I. 663.

Thomson Reuters once again moved for partial summary judgment on direct copyright infringement and related defenses. D.I. 674. Ross moved for summary judgment

As I later explain, I hold that:

- Ross infringed 2,243 headnotes. As to those headnotes, the only remaining factual issue is whether some of their copyrights have expired.
- Ross’s innocent infringement, copyright misuse, merger, and *scenes à faire* defenses all fail.

A. Direct copyright infringement

Thomson Reuters alleges that Ross directly infringed its copyrights. To show that, Thomson Reuters must show both that (1) it owned a valid copyright and (2) Ross copied protectable elements of the copyrighted work. *Feist*, 499 U.S. at 361. The second element requires showing both that (2a) Ross actually copied the work and that (2b) its copy was substantially similar to the work. *Dam Things from Den. v. Russ Berrie & Co.*, 290 F.3d 548, 561–62 (3d Cir. 2002).

(i) Copyright validity

Copyright validity is a question of law, not fact, making it suitable for summary judgment. *See Silvertop Assocs. Inc. v. Kangaroo Mfg. Inc.*, 931 F.3d 215, 219 (3d Cir. 2019). But an underlying factual dispute remains here. So the jury may need to decide this issue, but for reasons different from the ones I gave in my prior summary-judgment opinion.

Thomson Reuters must show that it had a valid copyright. Copyright registrations are “prima facie evidence of the validity of the copyright” if “made before or within five years after first publication of the work.” 17 U.S.C. § 410(c). Thomson Reuters has copyright registrations for Westlaw’s copyrightable content. D.I. 752-1 at 4. And it docketed registrations from 1981 to 2019. D.I. 1-1. So it has a valid compilation

whether a work is original, not how much effort went into developing it. *Id.* at 359–60. So I now revise those parts of my prior opinion. Now I see no genuine dispute that the headnotes and Key Number System clear *Feist*'s minimal threshold for originality.

1. *The headnotes are original.* A headnote is a short, key point of law chiseled out of a lengthy judicial opinion. The text of judicial opinions is not copyrightable. *Banks v. Manchester*, 128 U.S. 244, 253–54 (1888). And even if it were, Thomson Reuters would not get that copyright because it did not write the opinions. But a headnote can introduce creativity by distilling, synthesizing, or explaining part of an opinion, and thus be copyrightable. That is why I have changed my mind.

First, the headnotes are a compilation. “Factual compilations” are original if the compiler makes “choices as to selection and arrangement” using “a minimal degree of creativity.” *Feist*, 499 U.S. at 348. Thomson Reuters’s selection and arrangement of its headnotes easily clears that low bar.

More than that, each headnote is an individual, copyrightable work. That became clear to me once I analogized the lawyer’s editorial judgment to that of a sculptor. A block of raw marble, like a judicial opinion, is not copyrightable. Yet a sculptor creates a sculpture by choosing what to cut away and what to leave in place. That sculpture is copyrightable. 17 U.S.C. § 102(a)(5). So too, even a headnote taken verbatim from an opinion is a carefully chosen fraction of the whole. Identifying 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. That editorial expression has enough

“creative spark” to be original. *Feist*, 499 U.S. at 345. So all headnotes, even any that quote judicial opinions verbatim, have original value as individual works.

That belated insight explains my change of heart. In my 2023 opinion, I wrongly viewed the degree of overlap between the headnote text and the case opinion text as dispositive of originality. 694 F. Supp. 3d at 478. I no longer think that is so. But I am still not granting summary judgment on any headnotes that are verbatim copies of the case opinion (for reasons that I explain below).

2. *The Key Number System is original too.* There is no genuine issue of material fact about the Key Number System’s originality. Recall that Westlaw uses this taxonomy to organize its materials. Even if “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,” it still has the minimum “spark” of originality. *Id.* at 477 (internal quotation marks omitted); *Feist*, 499 U.S. at 345. The question is whether the system is original, not how hard Thomas Reuters worked to create it. *Feist*, 499 U.S. at 359–60. So whether a rote computer program did the work is not dispositive. And it does not matter if the Key Number System categorizes opinions into legal buckets that any first-year law student would recognize. To be original, a compilation need not be “novel,” just “independently created by” Thomson Reuters. *Id.* at 345–46. There are many possible, logical ways to organize legal topics by level of granularity. It is enough that Thomson Reuters chose a particular one.

Thus, I grant summary judgment for Thomson Reuters on whether the headnotes and the Key Number System are original enough to prevent Ross from rebutting any presumption of validity.

(ii) Copying of original elements

Next, I turn to whether there was “copying of constituent elements of the work that are original.” *Feist*, 499 U.S. at 361. I must decide whether Thomson Reuters has proven both (a) actual copying and (b) substantial similarity. *Dam Things*, 290 F.3d at 561–62.

1. *A few preliminaries.* Before applying that analysis, I must determine which parts of the case to apply it to. Ross moves for summary judgment on all the pieces that Thomson Reuters accuses Ross of infringing: 21,787 headnotes, the editorial decisions in 500 judicial opinions, and West’s Key Number System. D.I. 684 at 9–11. I see no proof sufficient to take all these items away from a jury.

Thomson Reuters, for its part, moves for summary judgment on only two subsets of the headnotes. D.I. 694 at 20 n.8. The two batches contain 5,367 and 2,830 headnotes each.

I reach no new decision on the Key Number System. There are still factual disputes about whether Ross used it and how, so I cannot analyze what protected elements Ross may have used. So I leave my prior ruling on this topic in place. Nor do I decide the fate of the 500 judicial opinions containing Thomson Reuters’s editorial decisions, because there are still factual questions about how and how much Ross accessed that material. For now, I consider only the batch of 2,830 headnotes identified by Thomson Reuters.

I leave the 5,367 for trial because Thomson Reuters's argument for summary judgment on that batch is flawed. Thomson Reuters points out that after I ordered Ross to submit a list of headnotes that are verbatim or near-verbatim quotations of judicial opinions, Ross did not submit these 5,367. D.I. 675 at 24. And it contends that Ross thereby admitted that these headnotes are protectable. *Id.* Perhaps I could view that as a concession about how similar the headnotes are to the judicial opinions. But Ross's omission makes no concession about the more important issue: whether the Bulk Memo questions copied these headnotes, or even whether the questions are closer to the headnote or to the judicial opinion. So it is not appropriate to grant summary judgment to Thomson Reuters on the 5,367 headnotes based on this reasoning. (Because I did not look at this batch, I need not address Ross's objection that it had no chance to respond to Thomson Reuters's expert's opinion on the 5,367 headnotes. D.I. 749.)

But it is appropriate to address the other 2,830 now at summary judgment. Before I explain why and how I address this batch, I must first clean up whether the batch in fact has 2,830 headnotes. There are two disputes. First, Ross's expert, Barbara Frederiksen-Cross, put 3,384 headnotes in this category, but the parties dispute whether 554 of them are still covered by valid copyright registrations. To resolve this, I looked at all 3,384 but make summary judgment contingent on the jury's findings about which 2,830 (or other number) still have valid copyrights.

Second, Ross claims that Thomson Reuters never asserted that 1,623 of the 2,830 headnotes were infringed, so at first, I denied summary judgment on them. 694

F. Supp. 3d at 479. Ross says I should again deny summary judgment because Thomson Reuters did not include these 1,623 headnotes when I ordered it to produce a list of allegedly infringing headnotes. D.I. 201; D.I. 324 ¶¶ 28–30; D.I. 324-23 (Ross’s spreadsheet identifying the 1,623). But on the same day that Thomson Reuters docketed its interrogatory response lacking those 1,623 headnotes, Ross docketed the expert report conceding them. D.I. 266-1 at 431 (Frederiksen-Cross’s report); D.I. 281-3–281-8 (Thomson Reuters’s interrogatory response). So Thomson Reuters did not have a chance to analyze Ross’s expert report before submitting its list of headnotes. Plus, Ross had the chance to respond to these 1,623 headnotes when Thomson Reuters raised them at summary judgment. And Ross’s own expert analyzed them, so there was no danger of unfair surprise. I revise my prior ruling and now consider these 1,623 headnotes as fairly part of the case.

Having sorted through that morass, I apply the actual-copying and substantial-similarity analyses to the 3,384 headnotes, which include the 2,830.

2. Actual copying. Actual copying means that “the defendant did, in fact, use the copyrighted work in creating his own.” *Tanksley v. Daniels*, 902 F.3d 165, 173 (3d Cir. 2018). One can prove this directly, with evidence that the defendant copied the work, or indirectly, by showing that the defendant had access to it and produced something similar (“probative similarity”). *Id.*

When evaluating copying, I may consider expert opinions. *See Kay Berry, Inc. v. Taylor Gifts, Inc.*, 421 F.3d 199, 208 (3d Cir. 2005). Frederiksen-Cross submitted an expert report stating that the Bulk Memo questions for this batch closely resemble

the headnotes' text and that the headnotes differ significantly from the text of the judicial opinions. D.I. 266-1 at 431, 435; D.I. 675 at 10. Her finding suggests that these questions were created by copying Westlaw headnotes, not by summarizing the underlying opinions. I view that as Ross's expert conceding that the 2,830 were actually copied.

But to make sure, the Court has now compared how similar each of the 2,830 Bulk Memo questions, headnotes, and judicial opinions are, one by one. The parties agree that LegalEase had access to Westlaw and used it to make the Bulk Memos. Of course, access alone is not proof. But a Bulk Memo question that looks more like a headnote than it does like the underlying judicial opinion is strong circumstantial evidence of actual copying. My comparison of the questions, headnotes, and opinions can decide whether probative similarity is present, so long as I think no reasonable jury could reach a different conclusion. Access plus probative similarity adds up to evidence of actual copying. To be sure, there was some confusion at the summary judgment hearing about whether Ross used all the Bulk Memos. *See* D.I. 769 at 30 (10:48:35), 162 (14:36:24). Ross's counsel said that some Bulk Memos were discarded but twice confirmed that Ross had used 80% for initial training and 20% for later validation. D.I. 769 at 30 (10:48:35), 162 (14:36:24). So taking counsel at his word, Ross used practically 100% to train its AI.

Having slogged through all 2,830 headnotes, I grant summary judgment to Thomson Reuters on actual copying, finding actual copying of 2,243. Appendix A to this opinion, filed under seal, catalogues the specific headnotes. I grant summary

judgment only on the headnotes for which actual copying is so obvious that no reasonable jury could find otherwise.

3. *Substantial similarity*. Substantial similarity requires evaluating whether “the later work materially appropriates the copyrighted work.” *Tanksley*, 902 F.3d at 173. That means deciding which parts of the actually copied work are original expression and so protected by copyright. Substantial similarity is often “an extremely close question of fact,” and thus well suited to the jury. *Id.* at 171 (internal quotation marks omitted). But summary judgment can be “appropriate” when “no reasonable jury could find” otherwise. *Id.* (internal quotation marks omitted). The question is whether an ordinary user of a product would find it substantially similar to the copyrighted work. *Dam Things*, 290 F.3d at 562. As a lawyer and judge, I am myself an ordinary user of Westlaw headnotes. So I am well positioned to determine substantial similarity here. I do so cautiously, and only on those headnotes for which I am confident that a reasonable jury could not conclude otherwise.

Ross argues that the Bulk Memo questions must be not just substantially similar to the headnotes, but virtually identical. The Ninth Circuit takes this approach to “thin” copyrights because “the range of protectable ... expression is narrow.” *Apple Comput., Inc. v. Microsoft Corp.*, 35 F.3d 1435, 1439, 1446 (9th Cir. 1994). But substantial similarity is always a spectrum, whether the copyright is “thin” or “thick.” “More similarity is required when less protectable matter is at issue.” 4 *Nimmer on Copyright* § 13D.32[A]. Other terms for this concept include Nimmer’s “supersubstantial similarity” and the Second Circuit’s “more discerning” ordinary-observer test. *Id.*;

Knitwaves, Inc. v. Lollytogs Ltd., 71 F.3d 996, 1002–03 (2d Cir. 1995). Though I do not apply the Ninth Circuit’s test explicitly, I do apply the concept underlying all these terms: The less protectable expression a work contains, the more similar the allegedly infringing work must be to it.

Applying this standard, I grant summary judgment on substantial similarity for Thomson Reuters on the 2,243 headnotes listed in Appendix A, finding that the Bulk Memo questions were substantially similar to them. Again, I grant summary judgment only on the headnotes for which substantial similarity is so obvious that no reasonable jury could find otherwise. In practice, this means that I am granting summary judgment only on the headnotes whose language very closely tracks the language of the Bulk Memo question but not the language of the case opinion. The rest of the headnotes must go to trial. I do not grant summary judgment to Ross on any headnotes because there are none for which I am confident that a reasonable jury could not find infringement. I do not decide at summary judgment the factual question of which headnotes are still covered by Thomson Reuters’s existing copyrights and leave this question open for trial.

B. Ross’s defenses to copyright infringement fail

None of Ross’s possible defenses holds water. I reject them all.

First, innocent infringement does not apply. Ross claims that any infringement was innocent. As the parties agree, innocent infringement does not limit liability, just damages. 17 U.S.C. § 504(c)(2). But this limit does not apply when the infringed work bears a copyright notice, as Westlaw’s headnotes do. 17 U.S.C. § 401(d).

Second, copyright misuse does not apply either. Ross claims that Thomson Reuters misused its copyright. Copyright misuse is a defense when a copyright holder weaponizes the copyright against the public interest, typically for “anti-competitive behavior.” *Video Pipeline, Inc. v. Buena Vista Home Ent., Inc.*, 342 F.3d 191, 203–06 (3d Cir. 2003). But as I already ruled at summary judgment on Ross’s antitrust counterclaims, Ross has not shown that Thomson Reuters misused its copyrights to stifle competition. D.I. 669.

Third, the merger defense is inapt. Ross claims that any ideas were so close to the expression that they merged with the expression, making it uncopyrightable. *Educ. Testing Servs. v. Katzman*, 793 F.2d 533, 539 (3d Cir. 1986). But there are many ways to express points of law from judicial opinions, so I reject this defense as well.

Fourth, the *scenes à faire* defense does not fit. This defense covers stock elements that follow from the work’s nature, like a historical romance novel’s damsel in distress. *Atari, Inc. v. N. Am. Philips Consumer Elecs. Corp.*, 672 F.2d 607, 616 (7th Cir. 1982). But nothing about a judicial opinion requires it to be slimmed down to Thomson Reuters’s headnotes or categorized by key numbers.

III. THOMSON REUTERS, NOT ROSS, PREVAILS ON THE FAIR-USE DEFENSE

There remains one more defense. In my 2023 opinion, I denied summary judgment on fair use. D.I. 548; 694 F. Supp. 3d at 482–87. But with new information and understanding, I vacate those sections of that order and its accompanying opinion addressing fair use. Fair use is an affirmative defense, so Ross bears the burden of proof. *Video Pipeline*, 342 F.3d at 197.

I must consider at least four fair-use factors: (1) the use’s purpose and character, including whether it is commercial or nonprofit; (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 Ross’s use affected the copyrighted work’s value or potential market. 17 U.S.C. § 107(1)–(4). The first and fourth factors weigh most heavily in the analysis. *Authors Guild v. Google, Inc.*, 804 F.3d 202, 220 (2d Cir. 2015) (Leval, J.).

“Fair use is a mixed question of law and fact.” *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 560 (1985). But “[i]n this case, the ultimate ‘fair use’ question primarily involves legal work.” *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 24 (2021). The undisputed facts here push this case squarely into the legal realm. Once we get past actual copying, the remaining issues that matter are not ones of historical fact, intent, or factual prediction. They are about how to apply the law to the facts. So here, fair use is a question for the judge, not the jury. Thomson Reuters prevails on the two most important and on the overall balancing.

A. Factor one goes to Thomson Reuters

First, I consider the purpose and character of Ross’s use. 17 U.S.C. § 107(1). I look mainly at whether it was commercial and whether it was transformative. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 529–31 (2023). If Ross and Thomson Reuters use copyrighted material like the headnotes for very similar purposes and Ross’s use is commercial, this factor likely disfavors fair use. *Id.* at 532–33.

1. *Ross’s use is commercial.* Ross admits as much. D.I. 727 at 29. It “stands to profit from exploitation of the copyrighted material without paying the customary

price.” *Harper & Row*, 471 U.S. at 562. But commerciality is not dispositive. I must balance it against how different this work’s purpose or character is. *Warhol*, 598 U.S. at 525.

2. *Ross’s use is not transformative.* Transformativeness is about 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*, 598 U.S. at 532–33. It weighs against fair use here. Ross’s use is not transformative because it does not have a “further purpose or different character” from Thomson Reuters’s. *Id.* at 529.

Ross was using Thomson Reuters’s headnotes as AI data to create a legal research tool to compete with Westlaw. It is undisputed that Ross’s AI is not generative AI (AI that writes new content itself). Rather, when a user enters a legal question, Ross spits back relevant judicial opinions that have already been written. D.I. 723 at 5. That process resembles how Westlaw uses headnotes and key numbers to return a list of cases with fitting headnotes. Thomson Reuters uses its headnotes and Key Number System primarily to help legal researchers navigate Westlaw and (possibly, as the parties dispute this) to improve Westlaw’s internal search tool. D.I. 769 at 14 (10:24:52). The parties agree that Ross and Westlaw are competitors. D.I. 752-1 at 4. So at first glance, this factor looks simple.

But, as Ross argues, the headnotes do not appear as part of the final product that Ross put forward to consumers. The copying occurred at an intermediate step: Ross

turned the headnotes into numerical data about the relationships among legal words to feed into its AI. D.I. 727 at 22. That makes this factor much trickier.

Ross is right that intermediate copying has been permitted under fair use factor one in analyzing computer programs. *See Google*, 593 U.S. at 30–32; *Sony Comput. Ent., Inc. v. Connectix Corp.*, 203 F.3d 596, 599, 606–07 (9th Cir. 2000) (holding that copying source code to create a product that lets people play Sony games on a personal computer, rather than a separate Sony game station, is transformative); *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1514–1515, 1522–23 (9th Cir. 1992) (holding that copying source code to create games that are compatible with an existing gaming system is transformative).

But those cases are inapt. First and foremost, those cases are all about copying computer code. This case is not. (Though Ross did computer programming, the material it allegedly copied from Thomson Reuters was not computer code.) In copyright, “computer programs differ from books, films, and many other literary works in that such programs almost always serve functional purposes.” *Google*, 593 U.S. at 21 (internal quotation marks omitted). So the fair-use considerations for these programs do not always apply to cases about copying written words.

Second and relatedly, these computer-programming cases about intermediate copying rely on a factor absent here: The copying was *necessary* for competitors to innovate. In *Google*, Google had copied part of a computer-programming language—specifically, the code that lets programmers speak to software in a particular way. *Id.* at 6, 29–33. That copying was “necessary for different programs to speak to each other.” *Id.* at

31. The copying in *Sony* was also necessary. The Ninth Circuit “appl[ie]d fair use to intermediate copying [that was] necessary to reverse engineer access to unprotected functional elements within a program.” *Id.* at 22. “[I]ntermediate copying ... was a fair use for the purpose of gaining access to the unprotected elements of Sony’s software.” *Sony*, 203 F.3d at 602. Likewise, *Sega* addressed copying that occurred “solely in order to discover the functional requirements for compatibility.” 977 F.2d at 1522. Here, though, there is no computer code whose underlying ideas can be reached only by copying their expression. The “copying is [not] reasonably necessary to achieve the user’s new purpose.” *Warhol*, 598 U.S. at 532.

My prior opinion wrongly concluded that I had to send this factor to a jury. 694 F. Supp. 3d at 483–84. I based that conclusion on *Sony* and *Sega*. Since then, I have realized that the intermediate-copying cases (1) are computer-programming copying cases; and (2) depend in part on the need to copy to reach the underlying ideas. Neither is true here. Because of that, this case fits more neatly into the newer framework advanced by *Warhol*. I thus look to the broad purpose and character of Ross’s use. Ross took the headnotes to make it easier to develop a competing legal research tool. So Ross’s use is not transformative. Because the AI landscape is changing rapidly, I note for readers that only non-generative AI is before me today.

3. *Even if relevant, bad faith would not move the needle.* The Supreme Court has expressed “some skepticism about whether bad faith has any role in a fair use analysis.” *Google*, 593 U.S. at 32. If it does, a reasonable jury might find that Ross, by going forward and arguably inducing LegalEase’s copying after Thomson Reuters refused

to license its content, acted in bad faith. But because Ross’s use was commercial and not transformative, I need not consider this possible element. Even if I found no bad faith, that finding would not outweigh the other two considerations.

B. Factor two goes to Ross

Second, I ask about the nature of the original work. 17 U.S.C. § 107(2). That involves “focus[ing] on the degree of creativity inherent to the work.” 4 *Nimmer on Copyright* § 13F.06. More creative works get more protection. *Id.* § 13F.06[A].

Westlaw’s material has more than the minimal spark of originality required for copyright validity. But the material is not *that* creative. Though the headnotes required editorial creativity and judgment, that creativity is less than that of a novelist or artist drafting a work from scratch. And the Key Number System is a factual compilation, so its creativity is limited.

I signaled a similar leaning before. 694 F. Supp. 3d at 484–85. Yet I stopped short of granting summary judgment based on factual disputes about how much creativity was involved. Now, as I concluded above, there is no factual dispute that the headnotes have creative elements but are far from the most creative works.

So factor two goes for Ross. Note, though, that this factor “has rarely played a significant role in the determination of a fair use dispute.” *Authors Guild*, 804 F.3d at 220.

C. Factor three goes to Ross

Third, I focus on how much of the work was used and how substantial a part it was relative to the whole. 17 U.S.C. § 107(3). I ask whether that usage was “reasonable in relation to the purpose of the copying.” *Campbell v. Acuff-Rose Music, Inc.*, 510

U.S. 569, 586 (1994). Courts consider both “the quantity of the materials used” and “their quality and importance.” *Id.* at 587. To win on this factor, the alleged copier must not take the “heart” of the work. *Id.*

My prior opinion did not decide factor three but suggested that it leaned towards Ross. The opinion focused on Ross’s claim that its output to an end user is a judicial opinion, not a West headnote, so it “communicates little sense of the original.” 649 F. Supp. 3d at 485 (quoting *Authors Guild*, 804 F.3d at 223).

I stand by that reasoning, but now go a step further and decide factor three for Ross. There is no factual dispute: Ross’s output to an end user does not include a West headnote. What matters is not “the amount and substantiality of the portion used *in making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a public for which it may serve as a competing substitute.” *Authors Guild*, 804 F.3d at 222 (internal quotation marks omitted). Because Ross did not make West headnotes available to the public, Ross benefits from factor three.

In its briefing, Ross emphasized that the number of headnotes allegedly taken amounted to only a small percentage of the total headnotes owned by Westlaw. That argument is inapt. If taking 300 words out of President Ford’s memoirs could count as taking the heart of the work, so too can taking several thousand headnotes out of Westlaw. *Campbell*, 510 U.S. at 587. The percentage of a total work copied is neither necessary nor sufficient to decide factor three. But Ross wins this factor anyway.

D. Factor four goes to Thomson Reuters

Factor four “is undoubtedly the single most important element of fair use.” *Harper & Row*, 471 U.S. at 566. For this factor, I consider the “likely effect [of Ross’s copying]

on the market for the original.” *Campbell*, 510 U.S. at 590. I must consider not only current markets but also potential derivative ones “that creators of original works would in general develop or license others to develop.” *Id.* at 592. I also consider any “public benefits the copying will likely produce.” *Google*, 593 U.S. at 35. The original market is obvious: legal-research platforms. And at least one potential derivative market is also obvious: data to train legal AIs.

My prior opinion left this factor for the jury. I thought that “Ross’s use might be transformative, creating a brand-new research platform that serves a different purpose than Westlaw.” 694 F. Supp. 3d at 486. If that were true, then Ross would not be a market substitute for Westlaw. Plus, I worried whether there was a relevant, genuine issue of material fact about whether Thomson Reuters would use its data to train AI tools or sell its headnotes as training data. *Id.* And I thought a jury ought to sort out “whether the public’s interest is better served by protecting a creator or a copier.” *Id.*

In hindsight, those concerns are unpersuasive. Even taking all facts in favor of Ross, it meant to compete with Westlaw by developing a market substitute. D.I. 752-1 at 4. And it does not matter whether Thomson Reuters has used the data to train its own legal search tools; the effect on a *potential* market for AI training data is enough. Ross bears the burden of proof. It has not put forward enough facts to show that these markets do not exist and would not be affected.

Nor does a possible benefit to the public save Ross. Yes, there is a public interest in accessing the law. But legal opinions are freely available, and “the public’s interest

in the subject matter” alone is not enough. *Harper & Row*, 471 U.S. at 569. The public has no right to Thomson Reuters’s parsing of the law. Copyrights encourage people to develop things that help society, like good legal-research tools. Their builders earn the right to be paid accordingly. This case is distinguishable from *Google*, where the API was valuable “because users, including programmers, [were] just used to it.” 593 U.S. at 38. 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.

E. Balancing the factors, I reject Ross’s fair-use defense

Factors one and four favor Thomson Reuters. Factors two and three favor Ross. Factor two matters less than the others, and factor four matters more. Weighing them all together, I grant summary judgment for Thomson Reuters on fair use.

I grant partial summary judgment to Thomson Reuters on direct copyright infringement for the headnotes in Appendix A. For those headnotes, the only remaining factual issue on liability is that some of those copyrights may have expired or been untimely created. This factual question underlying copyright validity is for the jury. I also grant summary judgment to Thomson Reuters against Ross’s defenses of innocent infringement, copyright misuse, merger, *scenes à faire*, and fair use. I deny Ross’s motions for summary judgment on direct copyright infringement and fair use. I revise all parts of my prior opinions that conflict with this one. I leave undisturbed the parts of my prior opinion not addressed in this one, such as my rulings on contributory liability, vicarious liability, and tortious interference with contract.

THOMPSON REUTERS CENTRE GMBH;

WEST PUBLISHING HOUSE

v.

ROSS INTELLIGENCE INC,

Petitioner

OFFICE OF THE CLERK

PATRICIA S. DODSZUWEIT

UNITED STATES COURT OF APPEALS

TELEPHONE

CLERK

FOR THE THIRD CIRCUIT
21400 UNITED STATES COURTHOUSE
601 MARKET STREET

215-597-2995

PHILADELPHIA, PA 19106-1790

Website: www.ca3.uscourts.gov



April 15, 2025

Jack B. Blumenfeld, Esq.
Morris Nichols Arsht & Tunnell
1201 N Market Street
16th Floor
P.O. Box 1347
Wilmington, DE 19899

Jacob S. Canter, Esq.
Crowell & Moring
3 Embarcadero Center
26th Floor
San Francisco, CA 94111

Yar R. Chaikovsky, Esq.
White & Case
3000 El Camino Real
5 Palo Alto Square
9th Floor
Palo Alto, CA 94306

Mark S. Davies, Esq.
White & Case
701 13th Street NW
Washington, DC 20005

Kayvan M. Ghaffari, Esq.
Pillsbury Winthrop Shaw & Pittman
Four Embarcadero Center
22nd Floor
San Francisco, CA 94111

Keith J. Harrison, Esq.
Crowell & Moring
1001 Pennsylvania Avenue NW
Washington, DC 20004

Andrew M. Legolvan, Esq.
White & Case
555 S Flower Street
Suite 2700
Los Angeles, CA 90071

David E. Moore, Esq.
Potter Anderson & Corroon
1313 N Market Street
6th Floor
Wilmington, DE 19801

Bindu A.G. Palapura, Esq.
Potter Anderson & Corroon
1313 N Market Street
6th Floor
Wilmington, DE 19801

Warrington S. Parker III, Esq.
Crowell & Moring
3 Embarcadero Center
26th Floor
San Francisco, CA 94111

Anne M. Voigts, Esq.
Pillsbury Winthrop Shaw Pittman
2250 Hanover Street
Palo Alto, CA 94304

RE: Thompson Reuters Centre GmbH, et al v. Ross Intelligence Inc
Case Number: 25-8018
District Court Case Number: 1:20-cv-00613

PACER account holders are required to promptly inform the PACER Service Center of any contact information changes. In order to not delay providing notice to attorneys or pro se public filers, your information, including address, phone number and/or email address, may have been updated in the Third Circuit database. Changes at the local level will not be reflected at PACER. Public filers are encouraged to review their information on file with PACER and update if necessary.

To All Parties:

The Clerk has received a petition for leave to appeal filed by **Ross Intelligence Inc**, docketed at **No. 25-8018**. All inquiries should be directed to your Case Manager in writing or by calling the

Clerk's Office at 215-597-2995. This Court's rules, forms, and case information are available on our website at <http://www.ca3.uscourts.gov>.

Counsel for Petitioner

As counsel for Petitioner(s), you must file:

1. Application for Admission (if applicable);
2. Appearance Form; and,
3. Disclosure Statement (except governmental entities), if not included in petition.

These forms must be filed within **fourteen (14) days of the date of this letter**.

Should the Court grant the petition for leave to appeal, additional forms and the appellate portion of the filing and docketing fee will be required.

Any response in opposition must be filed **within ten (10) days** of the date of this letter. **All responses must be accompanied by an Appearance Form and Disclosure Statement.** The petition and any response(s) will be forwarded to the Court for disposition. The parties will be advised when an order is entered by the Court.

Parties who do not intend to participate in the petition must notify the Court in writing.

Very truly yours,

Patricia S. Dodszeit, Clerk

By: Stephanie
Case Manager
Direct Dial 267-299-4926

cc: Honorable Stephanos Bibas
Randall C. Lohan, Esq.