IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

THOMSON REUTERS ENTERPRISE CENTRE GMBH and WEST PUBLISHING 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

on Thomson Reuters's copyright claims. D.I. 683. And both sides again moved for summary judgment on fair use. D.I. 672, 676. I now revise parts of my 2023 summary-judgment opinion. *See* Fed. R. Civ. P. 54(b).

I may grant summary judgment only if 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). I view all facts and draw all reasonable inferences in favor of the nonmovant. *Tundo v. County of Passaic*, 923 F.3d 283, 286–87 (3d Cir. 2019).

II. I GRANT PARTIAL SUMMARY JUDGMENT TO THOMSON REUTERS, NOT ROSS, ON DIRECT COPYRIGHT INFRINGEMENT AND RELATED DEFENSES

The dispute boils down to whether the LegalEase Bulk Memo questions copied Thomson Reuters's headnotes or were instead taken from uncopyrightable judicial opinions. To decide many issues here, one must compare the Bulk Memo questions, headnotes, and opinions side by side. I include the table below as an example. The questions and headnotes in this case are sealed. So the headnote and question in this table are not actual materials from the record, but an example I created based on *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345 (1991).

Question	West Headnote	Case Opinion
Does originality for copyright purposes mean that the work was independently created and has some minimal degree of creativity?	Originality, for copyright purposes, means that the work was independently created and has some minimal degree of creativity.	Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity.

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

copyright. But if Thomson Reuters chooses to try this case based on a theory of infringement of individual headnotes as individual works rather than infringement of the compilation as a whole, there is still a factual dispute about which individual headnotes are both within the period covered by Thomson Reuters's registrations and not in the public domain. *See* D.I. 755, 757, 761, 763. So, if Thomson Reuters advances a theory of damages that depends on the infringement of specific headnotes, this evidentiary matter must be taken up at trial.

In my 2023 opinion, I concluded that this issue would need to go to a jury, but for a different reason. I held that a jury would need to decide whether the headnotes and Key Number System were original enough. 694 F. Supp. 3d at 477–78.

Originality is central to copyright. The Constitution limits copyright protection to original works. *See Feist*, 499 U.S. at 346. So even if Thomson Reuters gets a presumption of validity because of the copyright registrations, Ross could rebut the presumption by showing that the works are not original. I previously thought that originality "depend[ed] on how much the headnotes overlap with the [uncopyrightable text of] opinions." 694 F. Supp. 3d at 478. And I explained that the Key Number System's originality was a jury question because Ross alleges that "most of the organization decisions are made by a rote computer program and the high-level topics largely track common doctrinal topics taught as law school courses." *Id.* at 477 (internal quotation marks omitted).

But the originality threshold is "extremely low," requiring only "some minimal degree of creativity …. some creative spark." *Id.* at 345. The key question, then, is

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[ied] 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.