25-2153

IN THE

United States Court of Appeals

FOR THE THIRD CIRCUIT

THOMSON REUTERS ENTERPRISE CENTRE GMBH
and WEST PUBLISHING CORPORATION,

Appellees,

against

ROSS INTELLIGENCE INC.,

Appellant.

On Appeal from an Order of
the United States District Court
for the District of Delaware Civil Action No. 20-613
(The Honorable Stephanos Bobas)

BRIEF FOR *AMICI CURIAE* BRIAN L. FRYE, JESS MIERS, AND MATEUSZ BLASZCZYK IN SUPPORT OF APPELLANT

Jeffrey Miles (#293869)
CARLTON FIELDS LLP
2029 Century Park East, Suite 1200
Los Angeles, California 90067
310-843-6300

Michael T. Hensley (#0391152001)
CARLTON FIELDS, P.A.
180 Park Avenue, Suite 106
Florham Park, New Jersey 07932
973-828-2600

Attroneys for Amici Curiae
Bryan L. Frye, Jess Miers, and Mateusz Blaszczyk

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

Rules

Other Authorities

RULE 29 STATEMENT

No party or party's counsel authored this brief in whole or in part or contributed money that was intended to fund preparing or submitting the brief. No person other than amici and their counsel contributed money that was intended to fund preparing or submitting the brief. See Fed. R. App. P. 29(a)(4)(E).

The undersigned law firm and its attorneys submit this brief solely as counsel of record for the amici law professors. In doing so, the law firm and its attorneys take no position on the substance of the arguments or opinions expressed herein. The views set forth in this brief are exclusively those of the amici law professors in their individual capacities and do not necessarily reflect the views of the law firm, its attorneys, or any of their clients.

INTEREST OF *AMICI CURIAE*

Brian L. Frye is the Spears-Gilbert Professor of Law at the University of Kentucky Rosenberg College of Law. As a copyright scholar, Professor Frye is interested in the scope and limits of United States copyright law. Accordingly, he believes it will be helpful to inform the court about the scope of copyrightable subject matter.

Jess Miers is a computer scientist and an Assistant Professor of Law at the University of Akron School of Law. Her scholarship focuses on the intersection of emerging technologies and expression. She urges courts to approach artificial intelligence with care, recognizing that new communications technologies can reshape public discourse for decades to come. Accordingly, she believes it will be helpful to inform the Court about the nuances of the training method at issue here.

Mateusz "Matt" Blaszczyk is a Research Fellow in Law and Mobility at the University of Michigan Law School. His scholarship focuses on intellectual property law and its intersections with AI. He believes that copyright doctrine in the area should be developed carefully and in accordance with principle. Accordingly, he believes it will be helpful to inform the court about recent developments in fair use theory, and how they might apply to AI training.

PRELIMINARY STATEMENT

While AI models present many difficult and important copyright questions, the most difficult and important question is existential: Does training an AI model on copyrighted works infringe the copyright in those works? Many federal courts are currently asking that question.[1] But this is the wrong case to answer it, because Thomson Reuters fails to allege that ROSS infringed any copyrightable elements of its work and fails to state a viable claim for copyright infringement.

The gravamen of Thomson Reuters's complaint is that ROSS infringed its copyright in West's headnotes and the West Key Number System by using them to train an AI model. Thomson Reuters fails to state a claim for copyright infringement for three reasons. First, West's headnotes consist entirely of uncopyrightable facts about the judicial opinions they describe. Second, the West Key Number System is in the public domain, because its copyright term has long since expired. And third, non-infringing uses of a work cannot be infringing, even if they affect the market for the work.

I. AI Training Takes Many Different Forms

Not every AI copyright infringement case is the same. Most of the headline cases involve the automated mass ingestion of copyrighted works. This case, by

[1] See generally Edward Lee, ChatGPT is Eating the World, at https://chatgptiseatingtheworld.com (reporting on the many lawsuits filed against AI companies alleging copyright infringement, among other things).

contrast, involves a distinct form of training that depends on independently created expert work product. The way the law is applied here could extend far beyond AI, threatening routine work like research and report writing. Even worse, it could wipe out the emerging job market of *human* experts hired to transfer their knowledge to AI in a manner compliant with existing copyright doctrine.

ROSS aims to make legal research more efficient by reducing the search costs imposed by traditional tools. Rather than sift through Boolean syntax and endless keyword hits, lawyers can pose their questions in plain English and receive direct excerpts from case law.

To teach its AI model, ROSS hired experts to create thousands of hand-crafted legal question-and-answer pairs. The ROSS experts engaged in their own legal research, using reputable legal databases, to create examples. These examples, later compiled into bulk training "memos," would teach the AI how to answer legal questions with relevant excerpts from on-point cases.

What ROSS's legal experts did to prepare the training memos was no different from what lawyers do when preparing briefs like this one. Yet it is this ordinary legal research task—and the expert-authored work it produced—that Thomson Reuters would paint as infringement with a broad brush, without pausing to make the threshold determination as to whether specific holdings expressed in each headnote that flowed into this process were themselves protectable.

A. Pretraining and Fine-Tuning Are Different

At its simplest, AI training proceeds in two distinct stages: pretraining and fine-tuning.[2] Both stages involve a plethora of different methods and training techniques, depending primarily on the goals of the AI's developers.[3] Of the two stages, pretraining is perhaps the most controversial, because it involves sweeping ingestion of data. As such, pretraining has become the focal point of most AI copyright litigation.[4]

1. Pretraining Uses Entire Works

The goal of pretraining is to give an AI model a foundational knowledge base. For language models—the backbone of today's chatbots—developers feed

---

[2] *See* Wayne Xin Zhao, et al., *A Survey of Large Language Models*, ARXIV:2303.18223 (Mar. 31, 2023), https://doi.org/10.48550/arXiv.2303.18223 (providing a comprehensive overview of the classic training paradigm for language models, including pretraining, fine-tuning, and alignment with human feedback).

[3] For example, pre-training and fine-tuning often involve a mix of supervised, unsupervised, and semi-supervised learning methods to enhance or activate model capabilities and align responses with human values. *Id.* at 14.

[4] *E.g.*, the New York Times alleges that OpenAI used millions of its articles to pretrain ChatGPT. *N.Y. Times Co. v. Microsoft Corp*., 777 F. Supp. 3d 283 (S.D.N.Y. 2025). Authors likewise claim that Meta relied on the controversial Books3 dataset—a corpora which included works ranging from self-published novels to books by Stephen King and Margaret Atwood—to pretrain its language models. *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC, 2025 WL 1752484 (N.D. Cal. June 25, 2025). And publishers identified their books among the materials Anthropic used to pretrain its Claude model, a case that settled after evidence showed Anthropic had drawn on pirated "shadow library" collections alongside lawfully acquired works. *Bartz v. Anthropic PBC*, No. C 24-05417 WHA, 2025 WL 1993577 (N.D. Cal. July 17, 2025).

them massive text collections, often drawn from books and Internet materials selected as reputable sources.[5] This exposure does not teach the model to memorize literal text, but to capture patterns of communication. As computing power grew and pretraining expanded to larger, more diverse datasets, researchers found that language models could generalize from those patterns to handle words, sentences, and questions they had never seen before. In other words, instead of simply remixing or parroting the pretraining texts, models began to create entirely new texts.[6] Essentially, pretraining produces a model that resembles a newly minted college graduate: able to converse across a wide range of topics, and—

[5] The data selected during pre-training depends on the goals and tasks of the AI system. Chatbots require immense amounts of text data. *See*, *e.g.*, Alex Reisner, *These 183,000 Books Are Fueling the Biggest Fight in Publishing and Tech*, THE ATLANTIC (Sept. 25, 2023), https://www.theatlantic.com/technology/archive/2023/09/books3-database-generative-ai-training-copyright-infringement/675363/ (describing the Books3 dataset, widely used in AI training and containing pirated book copies); Leo Gao, et al., *The Pile: An 800GB Dataset of Diverse Text for Language Modeling*, ARXIV:2101.00027 (Dec. 31, 2020), https://arxiv.org/abs/2101.00027 (introducing "the Pile," a widely used dataset for pretraining language models); *Common Crawl, About Common Crawl*, https://commoncrawl.org/overview (describing large-scale Internet text dataset obtained through web crawling); Xin Zhao, et al., *supra* note 2, at 14-15 (surveying large-scale pretraining datasets, including *Common Crawl*).

[6] *See* Jason Wei, et al., *Emergent Abilities of Large Language Models*, ARXIV:2206.07682 (last updated Oct. 26, 2022), https://arxiv.org/abs/2206.07682 (explaining that scaling up computing power and training data unlocks emergent capabilities in language models, including generalization—the ability to apply patterns learned during pretraining to novel tasks like question answering or summarization).

depending on the emphasis of their "studies" (the training data)—able to speak with greater depth in certain areas.[7]

Like most startup AI companies, ROSS did not construct its pretrained model from scratch. Its first product ran on IBM's Watson, an off-the-shelf natural language system already pretrained on massive, general-purpose text corpora. Eventually, ROSS moved away from Watson and began developing its own proprietary model. ROSS extended its base language model into the legal domain by adding judicial opinions to its pretraining corpus of text materials. This made the new pretrained base model much more capable of generalizing across the legal realm than Watson.

However, unlike in other AI copyright infringement cases, **pretraining is not at issue here.** Instead, this case concerns what ROSS did next: fine-tuning its model to specialize in answering legal research questions.

2. Fine-Tuning Uses Discrete Examples

Fine-tuning means giving a pretrained model specialized examples to teach it a new task—much like sending our proverbial college graduate to law school to

[7] Earlier language models trained on smaller, noisier datasets often did little more than memorize and remix their training text to mimic human conversation, earning the label "stochastic parrots." *See Emily M. Bender, et al., On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?,* ACM Conf. on Fairness, Accountability, and Transparency 610, *(2021) https://dl.acm.org/doi/pdf/10.1145/3442188.3445922.*

develop specialized expertise. With a broad pretrained foundation, a model can already generalize from just a few new examples, by learning rules and applying them to new problems.[8] And because fine-tuning requires specialized knowledge, developers typically rely on subject-matter experts to handcraft training examples.[9]

Legal education offers a useful parallel. Law students arrive "pretrained" by their undergraduate studies, already able to read, write, and reason. Law school then fine-tunes those skills for the legal domain. During their training, students do not need to memorize every case they read; instead, through hypotheticals and practice exams, they learn *patterns* of legal reasoning that allow them to tackle new questions on the bar exam. And behind that process is a professor who painstakingly designs the materials on which students train.

---

[8] *See* Xin Zhao, et al., *supra* note 2, at 14-15 (explaining that pretraining establishes broad linguistic ability, while fine-tuning aligns models with downstream tasks).

[9] The work of creating instruction–response pairs and labeling data for fine-tuning has given rise to a global "annotation" industry. Developers often contract with data-labeling vendors or hire annotators through online marketplaces. *See*, e.g., Amazon Mechanical Turk, https://www.mturk.com; Upwork, https://www.upwork.com; Fiverr, https://www.fiverr.com; *see also* r/mturk, Reddit, https://www.reddit.com/r/mturk/. These services typically rely on anonymous crowdworkers to perform low-cost microtasks. But annotation also extends to the high-skill, domain-expert end of the spectrum. LegalEase, for example, is a legal-process outsourcing firm staffed with licensed attorneys; rather than hiring anonymous annotators, ROSS contracted LegalEase to produce specialized, expert training data for its system.

Similarly, experts craft specialized training materials for AI—usually in the form of question–answer pairs or short instructions with expected responses.[10] In doing so, experts directly enhance pretrained models by conditioning them to produce reliable responses about specialized domains.[11]

One well-known example is OpenAI's GPT-3. Though fluent at generating new texts, it often ignored user instructions or generated unsafe responses.[12] OpenAI solved this problem by fine-tuning GPT-3 with thousands of expert-written instruction–response pairs like:[13]

---

[10] Examples of instruction-style pairs appear in the evaluation tasks used to test GPT-3. *See* Brown, *supra* note 6, app. G at 52 (illustrating a PIQA dataset prompt: "Context → My body cast a shadow over the grass because," with the correct answer "the sun was rising" and the incorrect answer "the grass was cut"). These examples show how language models can be conditioned at test time with "few-shot" prompts. In that setting, the model's weights are not updated; the examples only guide the response locally. By contrast, fine-tuning involves presenting the model with hundreds or thousands of such labeled examples, which update the model's weights and permanently embed the patterns into the system.

[11] Under the hood, each example adjusts the model's internal parameters, or weights, through repeated passes of gradient descent, a standard optimization method in neural networks. These weight updates "bake in" the patterns from the examples so that the model can apply its general language skills (that it learned during pretraining) reliably to new, unseen inputs. *See* Xin Zhao, et al., *supra* note 2, at 14-15.

[12] *See* OpenAI, *Aligning Language Models to Follow Instructions*, https://openai.com/index/instruction-following/ (Jan 27, 2022) (describing how GPT-3 models were fine-tuned with thousands of carefully engineered text prompts to improve their ability to perform specialized language tasks like question answering).

[13] *Id.*

- Instruction: "Explain the theory of gravity to a 6 year old."
- Response: "People went to the moon, and they took pictures of what they saw, and sent them back to the earth so we could all see them."

Instruction pairs like this one (multiplied by innumerable similar examples) trained GPT-3 to handle requests for simplification. Importantly, they do not ask the model to memorize or copy the example response; instead, they give the model a framework for appropriately responding to a particular kind of inquiry. Fine-tuning enabled OpenAI to transform GPT-3 into InstructGPT, the precursor to the revolutionary GPT-4.[14]

In other words, experts are indispensable to AI development. A developer building a model for mathematical proofs, for instance, might hire mathematicians to draft thousands of fine-tuning examples, consulting authoritative sources to ensure accuracy.[15] The need for fine-tuning has created a new market for domain

[14] *See* Long Ouyang, et al., *Training Language Models to Follow Instructions with Human Feedback*, ARXIV:2203.02155 (Mar. 4, 2022), at *2, https://arxiv.org/pdf/2203.02155 ("Specifically, we use reinforcement learning from human feedback to fine-tune GPT-3 to follow a broad class of written instructions").

[15] The work of creating instruction–response pairs and labeling data for fine-tuning has given rise to a global "annotation" industry. Developers often contract with data-labeling vendors or hire annotators through online marketplaces. *See*, e.g., Amazon Mechanical Turk, https://www.mturk.com; Upwork, https://www.upwork.com; Fiverr, https://www.fiverr.com; *see also* r/mturk, Reddit, https://www.reddit.com/r/mturk/. These services typically rely on anonymous crowdworkers to perform low-cost microtasks. But annotation also

specialists who can translate their expertise into training data—a market this case now threatens to erase by making anyone who hires experts to create such fine-tuning example an infringer.

B. How Did ROSS Fine-Tune Its AI Model?

With its pretrained foundation in place, ROSS hired legal experts from LegalEase Solutions to create thousands of question–answer pairs, producing a fine-tuning dataset known as the "Bulk Memos." This training corpus was designed to transform ROSS from a general-purpose, lawyerly *sounding* chatbot into a precise legal research engine.

Like the GPT-3 fine-tuning example, the ROSS memos were designed to model how the system should respond to legal research queries. Experts were instructed to draft realistic questions—sometimes guided, but not copied, from Westlaw headnotes—and pair them with excerpts from relevant judicial opinions.[16] To prepare these materials, they consulted reliable legal databases, selected case

---

extends to the high-skill, domain-expert end of the spectrum. LegalEase, for example, is a legal-process outsourcing firm staffed with licensed attorneys; rather than hiring anonymous annotators, ROSS contracted LegalEase to produce specialized, expert training data for its system.

[16] Def.'s Mot. Summ. J. Ex. 55, Whitehead Dep 127: 5-18, No. 690, *Thomson Reuters Enterprise Centre GMBH and West Publishing Corp. v. ROSS Intelligence, Inc.,* 1:20-cv-00613 (D. Del. Oct. 1, 2024) at *94, https://storage.courtlistener.com/recap/gov.uscourts.ded.72109/gov.uscourts.ded.72109.690.28.pdf.

passages they judged to be responsive, and labeled each excerpt as "Great" or "Good" (direct answers), "Topical" (related but not dispositive), or "Irrelevant" (superficially similar but off-point) so that the model could learn the difference between reliable and unreliable responses.

ROSS's experts did exactly what all legal professionals do when creating briefs or articles: consulting authoritative sources, exercising judgment, and framing original analysis. The LegalEase experts used headnotes as a research aid, then applied their expertise to draft questions and select passages from judicial opinions to teach the AI to determine whether law is dispositive, relevant, or irrelevant.

That fundamental fact places this case miles apart from current AI copyright suits premised on wholesale ingestion of copyrighted works. Unlike those cases, in which plaintiffs allege wholesale and indiscriminate copying of books or articles, the Bulk Memos were *newly created works*, crafted by experts from scratch, based upon public-domain judicial opinions and routine legal questions. And while the Westlaw Keynote System was indeed useful for this effort, Thomson Reuters does not have a monopoly over the practice of legal research.

C. Fine-Tuning Examples Are Not Infringing Reproductions

At bottom, Thomson Reuters's theory of infringement would discourage experts from consulting reference materials when creating new work, especially if

that work is then used to train AI. But no professional works in a vacuum: doctors rely on medical journals, engineers on technical manuals, and so on. Using Westlaw to identify issues and cases has always been part of ordinary legal practice.

But Thomson Reuters's theory of copyright infringement isn't just inconsistent with how lawyers use Westlaw and the myriad tools available, and necessary, to support effective legal research. It's also incompatible with copyright doctrine.

II. Thomson Reuters Fails to State a Claim for Copyright Infringement

In order to state a claim for copyright infringement, a plaintiff must allege both ownership of a valid copyright and actual copying of original elements of its copyrighted work. *See, e.g.*, *Feist Publications v. Rural Telephone Service Co.*, 499 U.S. 340, 361 (1991) ("To establish infringement, two elements must be proven: (1) ownership of a valid copyright, and (2) copying of constituent elements of the work that are original."). Thomson Reuters fails to state a claim for copyright infringement, because it cannot own a valid copyright in anything it alleges ROSS used to train its model. Thomson Reuters cannot own a copyright in West's headnotes, because they are uncopyrightable facts. And it cannot own a copyright in the West Key Number System, because it is in the public domain.

The district court found that ROSS copied West's individual headnotes, without reproducing any of the specific headnotes at issue nor specifying how it was that each of the specific headnotes differed from the underlying judicial opinions in a way that was unique and original. The district court found that ROSS copied a substantial collection of West's headnotes. That is irrelevant, because a mere collection of facts is only as copyrightable as the facts themselves—that is to say, without more, facts are not copyrightable. And the district court found that ROSS copied the West Key Number System. But that is irrelevant, because the West Key Number System is in the public domain.

A. West's Headnotes Aren't Copyrightable

The district court erred in finding that at least some of West's headnotes are copyrightable, because they aren't literally identical to the judicial opinions they describe. The district court was mistaken, because West's headnotes are intended to lack originality which is the *sine qua non* of copyrightability. The purpose of West's headnotes is not to be independently created works of authorship with a creative spark. The purpose of West's headnotes is to be accurate factual statements about judicial opinions. West's headnotes simply aren't the kind of thing that copyright protects.

1. West's Headnotes Aren't Original

Copyright can only protect original works of authorship. "The *sine qua non* of copyright is originality. To qualify for copyright protection, a work must be original to the author." *Feist*, 499 U.S. at 345 (citing *Harper & Row, Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 547-49 (1985)). A work or element of a work is original and protected by copyright if and only if it was "independently created" by the author of the work and reflects at least some degree of "creativity." *Feist*, 499 U.S. at 345 ("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.") (citing 1 M. Nimmer & D. Nimmer, Copyright §§ 2.01[A], [B] (1990)).

Of course, originality is a low bar. A work is independently created so long as it isn't a copy of an existing work, and originality requires only a "modicum" of creativity. *Feist*, 499 U.S. at 345-46 ("To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, 'no matter how crude, humble or obvious' it might be.") (quoting 1 M. Nimmer & D. Nimmer, Copyright § 1.08[C][1]). While copyright can't protect a white pages telephone directory, it can protect just about anything else, including a yellow pages telephone directory. But that doesn't help Thomson Reuters, because West's headnotes merely copy the

judicial opinions they describe and are intentionally devoid of creativity. Accordingly, copyright does not and cannot protect West's headnotes.

2. West's Headnotes Are Uncopyrightable Facts

Copyright cannot protect facts. "That there can be no valid copyright in facts is universally understood. The most fundamental axiom of copyright law is that '[n]o author may copyright his ideas or the facts he narrates.'" *Feist*, 499 U.S. at 344-45 (quoting *Harper & Row*, 471 U. S. at 556). The reason that copyright cannot protect facts is because facts are not independently created by an author and therefore cannot be original. "This is because facts do not owe their origin to an act of authorship. The distinction is one between creation and discovery: The first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence." *Feist*, 499 U.S. at 347.

In *Feist*, both parties agreed that copyright could not protect Rural's individual white pages telephone listings, because they were facts, and disagreed only about whether copyright could protect Rural's white pages telephone directory as a compilation of facts. *Feist*, 499 U.S. at 345 ("Rural wisely concedes this point, noting in its brief that '[f]acts and discoveries, of course, are not themselves subject to copyright protection.'"). According to the Supreme Court, Rural's white pages telephone listings were facts because they merely copied the name, address, and telephone number of the telephone customer they described.

Feist, 499 U.S. at 347 ("Census takers, for example, do not 'create' the population figures that emerge from their efforts; in a sense, they copy these figures from the world around them.").

West's headnotes are also facts, because they merely copy the judicial opinions they describe. Many of West's headnotes are literal copies of the judicial opinions they describe, consisting of a quotation. While some of West's headnotes are not literal copies of the opinions they describe, they are still mere paraphrases, intended to state a fact about the content of the opinion.

Once upon a time, Michelangelo supposedly quipped, "I saw the angel in the marble and carved until I set him free." The district court likens West's editors to Michelangelo, claiming that they carve headnotes from the raw stone of judicial opinions. *See Thomson Reuters Enterprise Center GMBH v. ROSS Intelligence Inc.*, 1:20-cv-613-SB (D. Del. Feb. 11, 2025) at 7. But West is in the business of condensing judicial opinions for the benefit of busy lawyers, not carving a *Pietà*.

While courts have held that legal reporters like West can own a copyright in the materials they create, they have never held that the individual headnotes in a legal digest are independently copyrightable. *See, e.g.*, *Georgia v. Public.Resource.Org, Inc.*, 140 S. Ct. 1498, 1507 (2020) and *Callaghan v. Myers*, 128 U.S. 617 (1888). Compare *West Pub. Co. v. Mead Data Cent., Inc.*, 799 F.2d 1219 (8th Cir. 1986) (holding before *Feist* that West's pagination was

copyrightable) with *Matthew Bender Co., Inc. v. West Publishing Co.*, 158 F.3d 693 (2d Cir. 1998) (holding after *Feist* that West's pagination system was not copyrightable). For good reason. The purpose of West's headnotes is to be accurate, not to be original. Indeed, if a West headnote were original, it would be defective. A West headnote is not supposed to be independently created, it is supposed to be a copy of the judicial opinion it describes. And a West headnote certainly isn't supposed to be creative, because any creativity would defeat its purpose. *See Southco, Inc. v. Kanebridge Corp.*, 390 F.3d 276, 282 (3d Cir. 2004) (en banc) (Alito, J.) ("Indeed, if any creativity were allowed to creep into the numbering process, the system would be defeated.").

The district court found that "even a headnote taken verbatim from an opinion" can be copyrightable since it is "a carefully chosen fraction of the whole." *See Thomson Reuters Enterprise Center GMBH v. ROSS Intelligence Inc.*, 1:20-cv-613-SB (D. Del. Feb. 11, 2025) at 3. But this would allow Thomson Reuters to own a copyright in a headnote—a "short, key point of law chiseled out of a lengthy judicial opinion"—even if it contains no new expression whatsoever, but is merely a quotation. *Id*. That would conflict with basic principles of copyrightability. *See*, *e.g.*, *Stewart v. Abend*, 495 U.S. 207, 234 (1990) (holding that an author "may receive protection only for his original additions," not "elements ... already in the public domain") and *ABS Ent., Inc. v. CBS Corp.*, 908 F.3d 405, 414 (9th Cir.

2018) ("A copy…is not a separate work, but a mere representation or duplication of a prior creative expression.").

Allowing Thomson Reuters to copyright the law simply by quoting it would privatize judicial opinions and statutes, undermining not only the public domain but also due process. *See, e.g.*, *Building Officials and Code Adm. v. Code Technology, Inc*., 628 F.2d 730, 734 (1st Cir. 1980) ("The citizens are the authors of the law, and therefore its owners, regardless of who actually drafts the provisions, because the law derives its authority from the consent of the public, expressed through the democratic process….This policy is, at bottom, based on the concept of due process.") and *Georgia v. Public.Resource.Org, Inc.*, 140 S. Ct. 1498, 1507 (2020) ("[N]o one can own the law."). *See also* U.S. Copyright Office, Compendium of U.S. Copyright Office Practices § 313.6(C)(2) (3d ed. 2021).

B. The West Key Number System is in the Public Domain

While it's unclear exactly when the West Key Number System was actually created, because it was based substantially on earlier systems of law digesting, the West Publishing Company began marketing the West Key Number System in 1909, marking the latest possible date of its creation and publication.

As such, while the district court assumed that the West Key Number System is protected by copyright, even that were true, it has been in the public domain for

decades. While the West Key Number System was probably protected by copyright when it was created, its copyright term expired at least two decades ago.

As the district court correctly observes, while copyright does not and cannot protect facts, it can and does protect compilations of facts. *See Thomson Reuters Enterprise Center GMBH v. ROSS Intelligence Inc*., 1:20-cv-613-SB (D. Del. Feb. 11, 2025) at 8. Or rather, copyright can and does protect the original elements of a compilation of facts, just like it protects the original elements of any other work. *See Feist Publications, Inc. v. Rural Tel. Serv. Co*., 499 U.S. 340, 357 (1991) (observing that "a compilation, like any other work, is copyrightable only if it satisfies the originality requirement"). Specifically, copyright can and does protect an original "selection, coordination, or arrangement" of facts in a compilation. *Id.*

The West Key Number System was probably copyrightable when it was created, because it comprises an original selection, coordination, or arrangement of facts about judicial opinions. *Thomson Reuters Enterprise Center GMBH v. ROSS Intelligence Inc*., 1:20-cv-613-SB (D. Del. Feb. 11, 2025) at 5 ("Thomson Reuters's selection and arrangement of its headnotes easily clears that low bar.") *See also* Callaghan v. Myers 128 U.S. 617 (1888) (assuming that a law digest includes copyrightable elements).

As of January 1, 2025, works published in the United States before 1930 are in the public domain. *See* 17 U.S.C. § 304(b) (creating a copyright term of 95 years

for works in their renewal term). The West Key Number System was published in 1909, so it entered the public domain in 2005 at the latest. Accordingly, it is currently in the public domain, and was also in the public domain in 2014 when ROSS allegedly infringed it.

The district court observes that Thomson Reuters owns many valid copyright registrations for Westlaw's copyrightable content. *See Thomson Reuters Enterprise Center GMBH v. ROSS Intelligence Inc*., 1:20-cv-613-SB (D. Del. Feb. 11, 2025) at 5-6. That is not the question at hand. The question before the court is not whether the Westlaw database contains any copyrightable content—it does—but whether ROSS copied any copyrightable elements of that database. It didn't.

Of course, Thomson Reuters owns a copyright in the new original written material created by West editors, including the syllabi they write for judicial opinions. *See Feist*, 499 U.S. 340 at 348 ("Thus, if the compilation author clothes facts with an original collocation of words, he or she may be able to claim a copyright in this written expression. Others may copy the underlying facts from the publication, but not the precise words used to present them.") But it no longer owns a copyright in the West Key Number System.

III. Dilution Is Not Copyright Infringement

Finally, the district court erred in its suggestion that Thomson Reuters can state a viable claim for copyright infringement based on ROSS's dilution of the value

of its copyrights. Dilution is not and cannot be a viable theory of copyright infringement.

In May 2025, the U.S. Copyright Office released a pre-publication report on generative AI training.[17] Among other things, the report observed that an AI model may generate content that is not substantially similar to any of the works in its training data but nevertheless dilutes the market for those works, and suggested that this kind of "dilution" could provide the basis for an infringement claim.[18] Acknowledging this to be "uncharted territory," the Office proposed to read the fourth factor of fair use as encompassing "any economic effect," including competing with any or all works of the same kind, regardless of similarities with particular works; or involving similarities in ideas or style, traditionally disregarded by the copyright doctrine.[19]

In *Kadrey v. Meta*, No. 23-cv-03417-VC, 2025 WL 1752484 at *16 (N.D. Cal. June 25, 2025), U.S. District Court Judge Vince Chhabria endorsed the Copyright Office's dilution theory of infringement. While the plaintiffs did not plead a dilution theory, Judge Chhabria encouraged them to do so, explaining in obiter dicta why the

[17] U.S. Copyright Office, Copyright and Artificial Intelligence Part 3: Generative AI Training. Pre-publication Version (May 2025), https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf [hereinafter *Report*].

[18] *Id.* at 73.

[19] *Id.* at 65-66.

Court would have found it compelling. However, in its issuance of the theory as dicta, the Court failed to substantiate dilution as a basis of a legitimate copyright infringement claim. According to Judge Chhabria, this is a theory of "indirect substitution," allowing a finding of infringement in the generative AI context, where "rapid generation of countless works that compete with the originals, even if those works aren't themselves infringing" is observed. *Id*. In other words, finding non-infringing works to be infringing simply because they are competitive in quality and cost of production. Notably, the court didn't provide any precedential authority to support the application of the dilutive theory.[20]

The dilution theory essentially holds that using an AI model to generate non-infringing content similar to its training data is a form of unfair competition. In lieu of copyright precedent, the Office's report cited with approval submissions from the Federal Trade Commission. The FTC proposed a similarly unorthodox solution: to use unfair competition law, specifically Section 5 of the FTC Act, to find infringement in scenarios which copyright doctrine considers fair use.[21] This

[20] *See Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC, 2025 WL 1752484, at *18 (refusing to "robotically apply[] concepts from previous cases").

[21] Report, *supra* note 26, at 75; *see also* Federal Trade Commission, Artificial Intelligence and Copyright Comment (No. 2023-6, Oct. 30, 2023), available at https://www.ftc.gov/system/files/ftc_gov/pdf/p241200_ftc_comment_to_copyright_office.pdf.

proposal has been criticized by scholars,[22] and has seemingly been abandoned since its cryptic issuance in the Office's report.[23] Most importantly, it is explicitly rooted in the assumption that fair use protects kinds of competition the FTC wanted to prohibit.

The problem with the dilution theory is that producing similar, but non-infringing works is precisely the kind of competition copyright is supposed to promote. In *Kadrey*, the district court explicitly opined that market dilution is "not the kind of competitive or creative displacement that concerns the Copyright Act." *Kadrey*, 2025 WL 1752484 at *2. The Act seeks to promote the creation of original works of authorship, not to protect authors against competition. Indeed, it is axiomatic that the purpose of copyright is to benefit the public by encouraging marginal authors to produce and distribute additional works of authorship.[24] It is thus

[22] *See e.g.*, Daryl Lim & Peter K. Yu, *The Antitrust-Copyright Interface in the Age of Generative Artificial Intelligence*, 74 EMORY L. J. 847 (2025).

[23] Andrew Ferguson, *Heeding The Rallying Cry for Deregulation*, Prepared Remarks at the International Competition Network Annual Conference 2025 Edinburgh, United Kingdom May 7, 2025, at *6, https://www.ftc.gov/system/files/ftc_gov/pdf/chairman-ferguson-2025-icn-remarks.pdf.

[24] *See e.g.*, *Fox Film Corp. v. Doyal*, 286 U.S. 123, 127 (1932) ("The sole interest of the United States and the primary object in conferring the monopoly lie in the general benefits derived by the public from the labors of authors."); United States v. Paramount Pictures, Inc., 334 U.S. 131, 158 (1948) ("The copyright law, like the patent statutes, makes reward to the owner a secondary consideration."); *Mazer v. Stein*, 347 U.S. 201, 219 (1954) ("The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that

unsurprising that Patry's treatise has dubbed the dilution theory an "erroneous" theory unsupported by any precedent, contrary to the statute and not based "on there being any infringement at the output stage."[25]

What's more, a regime based on dilutive infringement theory would be unconstitutional, because it allows a plaintiff to claim infringement based on non-infringing works.[26] According to the Supreme Court, "originality is a constitutionally mandated prerequisite for copyright protection." *Feist*, 499 U.S. at 351. But under the dilution theory, a plaintiff can show market harm caused by similar, but non-infringing AI-generated content. This would effectively allow copyright owners to claim copyright ownership of uncopyrightable elements of their works, which the Intellectual Property Clause disallows. *See Eldred v. Ashcroft*, 537 U.S. 186, 219 (2003) (observing that "every idea, theory, and fact in a copyrighted

---

encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'"); *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984) ("It is intended to motivate the creative activity of authors and inventors by the provision of a special reward, and to allow the public access to the products of their genius after the limited period of exclusive control has expired."); *Eldred v. Ashcroft*, 537 U.S. 186, 214 (2003) ("We can demur to petitioners' description of the Copyright Clause as a grant of legislative authority empowering Congress to secure a bargain—this for that.") (internal quotation marks omitted).

[25] *Market dilution theory*, *in* 4 PATRY ON COPYRIGHT § 10:155.40.

[26] *See id.* ("[T]he theory's entire purpose is to eliminate the 300-year old requirement in Anglo-American case law that harm to the market must be harm to the individual work and cannot be speculative.").

work becomes instantly available for public exploitation at the moment of publication").

There's no infringement without an infringing use. *See, e.g.*, *Baker v. Selden*, 101 U.S. 99, 104 (1879) ("The copyright of a book on bookkeeping cannot secure the exclusive right to make, sell, and use account books prepared upon the plan set forth in such book."). If ROSS didn't use any original elements belonging to Thomson Reuters, then it can't be a copyright infringer.

Nor would the theory espoused by the *Kadrey* Court apply to the instant case, since it is only in an exceptional generative AI circumstance that Judge Chhabria and the Office would consider it as a theory of infringement.[27] As the district court in this case recognized, ROSS's AI model isn't a generative AI model.[28] Accordingly, the dilution theory would not apply to ROSS even if it were a viable theory of infringement.

---

[27] *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC, 2025 WL 1752484, at *18 (N.D. Cal. June 25, 2025) ("It's true that, in many copyright cases, this concept of market dilution or indirect substitution is not particularly important…This case is different…[it] involves a technology that can generate literally millions of secondary works, with a miniscule fraction of the time and creativity used to create the original works it was trained on.").

[28] *Thomson Reuters Enter. Centre GmbH v. Ross Intell. Inc*., 765 F. Supp. 3d 382, 398 (D. Del. 2025) ("Ross's AI is not generative AI (AI that writes new content itself)"); *see also* Bartz v. Anthropic PBC, No. C 24-05417 WHA, 2025 WL 1741691, at *8 (N.D. Cal. June 23, 2025).

CONCLUSION

For the reasons stated above, this Court should reverse the district court's decision and remand the case for further proceedings.

DATED: September 29, 2025

Respectfully submitted,
/s/ Michael T. Hensley

Michael T. Hensley
CARLTON FIELDS, P.A.
Michael T. Hensley (# 0391152001)
180 Park Avenue, Suite 106
Florham Park, New Jersey 07932
Tel: (973) 828-2600

CARLTON FIELDS LLP
Jeffrey John Miles (# 293869)
2029 Century Park East, Suite 1200
Los Angeles, California 90067-2913
Tel: (310) 843-6300

Counsel for Amici Curae

COMBINED CERTIFICATIONS

STATEMENT OF BAR MEMBERSHIP

I, Michael T. Hensley, am duly admitted to practice law before the United States Court of Appeals for the Third Circuit.

/s/ Michael T. Hensley
Michael T. Hensley

STATEMENT OF COMPLIANCE WITH FRAP 32(a)(7)(b)

This brief was prepared on a computer using Microsoft Word. The font is Times New Roman, 14-point, double spaced. The word count of the body of the brief, including footnotes and point headings, as calculated by Microsoft Word, is 6,251 words.

/s/ Michael T. Hensley
Michael T. Hensley

STATEMENT OF SERVICE

I, Michael T. Hensley, hereby certify that I filed the foregoing document with the Clerk of the Court, using the CM/ECF system, which will automatically send notification and a copy of the brief to counsel who have appeared for the parties and are CM/ECF participants.

/s/ Michael T. Hensley
Michael T. Hensley

STATEMENT OF IDENTICAL COMPLIANCE BRIEFS

I, Michael T. Hensley, hereby certify that the briefs are served in compliance with the Court's rules and identical to one another and the electronic PDF version.

/s/ Michael T. Hensley
Michael T. Hensley

STATEMENT OF VIRUS CHECK

I, Michael T. Henlsey, hereby certify that a virus check was performed on this brief and that it is free from viruses. The check was performed with SentinelOne, version 25.1.3.334.

/s/ Michael T. Hensley
Michael T. Hensley