

No. 25-2153

United States Court of Appeals

for the

Third Circuit

THOMSON REUTERS ENTERPRISE CENTRE GMBH and WEST PUBLISHING
CORPORATION,

Plaintiff-Appellees,

v.

ROSS INTELLIGENCE INC.

Defendant-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 BIBAS)

**BRIEF FOR *AMICUS CURAE* FOUNDATION FOR AMERICAN
INNOVATION IN SUPPORT OF DEFENDANT-APPELLANT**

Foundation for American Innovation

Tim Hwang

2443 Fillmore St #380-3386

San Francisco, CA, 94115

(973)-960-4955

Attorney for Amicus Curae

CORPORATE DISCLOSURE STATEMENT

The Foundation for American Innovation (FAI) is a nonprofit corporation organized under Section 501(c)(3) of the Internal Revenue Code. FAI has no parent company and issues no stock. No company owns an interest of ten percent or more in FAI.

TABLE OF CONTENTS

	Page
TABLE OF AUTHORITIES	1
IDENTITY AND INTEREST OF AMICUS CURIAE	3
INTRODUCTION AND SUMMARY OF ARGUMENT	4
ARGUMENT	6
I. ONLY THIN COPYRIGHT PROTECTION EXISTS; ANY OVERLAP IS IN NONPUBLIC TRAINING, NOT IN USER OUTPUTS	6
II. FACTOR ONE: NONPUBLIC, FUNCTIONAL TRAINING CAN BE JUSTIFIED IN THIS CASE	8
A. The Statute And Case Law Show That Intermediate, Nonpublic Copying To Reach Functionality Can Satisfy Factor One Even In Commercial Settings	8
B. The Record Shows Nonpublic, Necessary Training That Extracts Functional Signals And Reveals No Headnote Prose, So Factor One Favors Fair Use	12
C. Medium Labels Are Irrelevant; What Matters Is The Necessity Of Intermediate Copying To Reach Functionality	15
III. FACTOR FOUR SAYS TO COUNT HARM ONLY TO MARKETS FOR PROTECTED EXPRESSION; NONE IS SHOWN HERE	18

A. The Relevant Market Is Licensing Of Headnote Expression—Not The Platform-level Market For Legal Research	18
B. Courts Find Harm Through Substitution, Usurpation Of Expressive Licensing Markets, Or Erosion Of Price or Exclusivity; None Is Present Here	20
C. Non-cognizable Theories Including “Training-license” Markets And “Market Dilution” Should Be Rejected	23
D. Policy Note	24
CONCLUSION	26
ADDENDUM: CONSENT OF THE PARTIES (FRAP 29(a)(2))	28
CERTIFICATE OF COMPLIANCE	
CERTIFICATE OF SERVICE	

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Am. Geophysical Union v. Texaco Inc.</i> , 60 F.3d 913 (2d Cir. 1995)	20, 21, 24
<i>Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith</i> , 598 U.S. 508 (2023)	Passim
<i>Authors Guild v. Google, Inc.</i> , 804 F.3d 202 (2d Cir. 2015)	Passim
<i>Authors Guild v. HathiTrust</i> , 755 F.3d 87 (2d Cir. 2014)	Passim
<i>Baker v. Selden</i> , 101 U.S. 99 (1880)	18, 19
<i>Banks v. Manchester</i> , 128 U.S. 244 (1888)	7, 23
<i>Bill Graham Archives v. Dorling Kindersley Ltd.</i> , 448 F.3d 605 (2d Cir. 2006)	19, 21, 23
<i>Campbell v. Acuff-Rose Music, Inc.</i> , 510 U.S. 569 (1994)	8, 15, 20, 21
<i>Castle Rock Entm’t, Inc. v. Carol Publ’g Grp., Inc.</i> , 150 F.3d 132 (2d Cir. 1998)	20
<i>Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.</i> , 499 U.S. 340 (1991)	18, 19, 23

<i>Fox News Network, LLC v. TVEyes, Inc.</i> , 883 F.3d 169 (2d Cir. 2018)	10, 20
<i>Google LLC v. Oracle Am., Inc.</i> , 593 U.S. 1 (2021)	Passim
<i>Harper & Row, Publ’rs, Inc. v. Nation Enters.</i> , 471 U.S. 539 (1985)	20, 22
<i>Kadrey v. Meta Platforms, Inc.</i> , No. 23-cv-03417-VC (N.D. Cal. June 25, 2025)	23, 24
<i>Kelly v. Arriba Soft Corp.</i> , 336 F.3d 811 (9th Cir. 2003)	20
<i>Perfect 10, Inc. v. Amazon.com, Inc.</i> , 508 F.3d 1146 (9th Cir. 2007)	19, 21
<i>Perris v. Hexamer</i> , 99 U.S. 674 (1879)	12
<i>Princeton Univ. Press v. Mich. Document Servs., Inc.</i> , 99 F.3d 1381 (6th Cir. 1996)	20, 21, 24
<i>Salinger v. Random House, Inc.</i> , 811 F.2d 90 (2d Cir. 1987)	20, 22
<i>Sega Enters. Ltd. v. Accolade, Inc.</i> , 977 F.2d 1510 (9th Cir. 1992)	Passim
<i>Sony Computer Entm’t, Inc. v. Connectix Corp.</i> , 203 F.3d 596 (9th Cir. 2000)	Passim
<i>Sony Corp. of Am. v. Universal City Studios, Inc.</i> , 464 U.S. 417 (1984)	16

<i>Video Pipeline, Inc. v. Buena Vista Home Entm't, Inc.</i> , 342 F.3d 191 (3d Cir. 2003)	10, 22
---	--------

Statutes

17 U.S.C. § 107(4)	Passim
17 U.S.C. § 102(b)	Passim
17 U.S.C. § 103(b)	5, 7, 18

IDENTITY AND INTEREST OF AMICUS CURIAE¹

The Foundation for American Innovation (FAI) is a 501(c)(3) nonprofit organization. Its mission is to develop technology, talent, and ideas that support a better, freer, and more abundant future. FAI advances this mission by helping policymakers foster innovation, encouraging technologists and founders to engage in governance, and informing the public about how such collaboration is vital to the continued success of the American project.

FAI files this brief because the question presented—how fair use applies when developers make nonpublic, intermediate copies to train an AI system—will set important precedent for AI, search, and accessibility tools. As the first appellate

¹ No party's counsel authored any part of this brief. No one, apart from FAI and its counsel, contributed money intended to fund this brief's preparation or submission. Special thanks to Joshua Levine and Samuel Roland for their help in putting together this brief.

case to address this issue, the Court’s decision will shape the legal environment for startups, researchers, and established firms building lawful, useful systems.

FAI has a strong institutional interest in clear fair-use rules for AI training and in keeping market-effect analysis confined to protected expression.

INTRODUCTION AND SUMMARY OF ARGUMENT

This case presents a narrow question with broad consequences: when a developer copies text at a nonpublic, intermediate stage to train an AI model, and the user-facing outputs are functional content (here, judicial opinions)—not the copied editorial prose—how should fair use apply? Though the lower court attempted to differentiate this case from new generative AI models, stating explicitly “Ross’s AI is not generative AI,” the analysis the court conducts belies the attempt.² Mem. Op. at 17–18. Nearly all AI models, whether search, text, vision, or other, undertake the same process of intermediate training and public output. Therefore, the reasoning of this Court—particularly as the first appeals circuit to rule on this issue—will establish an important precedent on all AI fair use. Below, we have provided our view on the most faithful reading of the applicable law in this novel context.

² *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intelligence Inc.*, No. 1:20-cv-613-SB (D. Del. Feb. 11, 2025) (Mem. Op.); ROSS Opening Br., D.I. 698 (D. Del. Oct. 1, 2024) (hereinafter “ROSS Opening Br.”).

The rule emerging from statute and precedent is clear: factor one favors fair use when **(1)** *the use of the copy at issue is an intermediate step to achieve an end distinct from the plaintiff's original expressive use*, and **(2)** *the taking is no more than reasonably necessary to reach the underlying functional ideas or interfaces*.

The rule has clear limits that protect the interests of the rightsholder. These limits preserve incentives for expressive works while supporting the continued progress in AI methods that is key to our nation's economic vitality. It does not excuse training or deployment that stores or outputs the plaintiff's prose (or derivative summaries) to end users; it does not excuse copying beyond what is reasonably necessary where less expressive alternatives would achieve comparable function; and it does not bless downstream products that embed headnote text or replicate West's selection/arrangement for public display. Those uses would defeat the presumption and create expressive substitution cognizable under 17 U.S.C. § 107(4).

The statute draws two lines. Section 107 separates purpose/character from market effect; 17 U.S.C. §§ 102(b), 103(b) limit protection to original expression, not ideas, facts, methods of operation, systems, or law/taxonomy. Any rule must keep factors one and four distinct. Precedent then supplies the standard. *Warhol* centers the “use at issue” and presumes factor one weighs against fair use when purposes are the same and the use is commercial—unless there is “other

justification.” *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 532–33 (2023). That justification is the settled doctrine permitting nonpublic, proportional intermediate copying to reach functional elements. The justification is derived from a line of cases dealing with functional learning rather than public communication, thereby supplying *Warhol’s* “other justification.” Factor four then asks only about markets for protected expression and traditional derivatives, not methods or access to public-domain law. Because outputs are opinions only, the relevant market is licensing of or replacement headnote text, and a bare nonpublic “license to be trained on” is not a market for expression; no cognizable harm appears on this record.

This Court should adopt the rule above and evaluate the record accordingly: treat nonpublic, reasonably necessary functional training as satisfying factor one and confine factor-four markets to expression and traditional derivatives.

ARGUMENT

I. ONLY THIN COPYRIGHT PROTECTION EXISTS; ANY OVERLAP IS IN NONPUBLIC TRAINING, NOT IN USER OUTPUTS

Only thin copyright protection remains in plaintiffs’ headnotes; any overlap, if it exists, is confined to nonpublic, intermediate training. The user-facing outputs are judicial opinions, not protected headnote prose. *Mem. Op.* at 21. Because only

headnote expression is protectable—and the outputs contain none of it—any alleged taking can occur, if at all, only during nonpublic training.

The analysis tracks the abstraction–filtration–comparison framework. Originality for headnotes and the Key Number System is assumed. Before any similarity assessment, the law excludes: (1) judicial opinions (*Banks v. Manchester*, 128 U.S. 244 (1888)); (2) ideas, procedures, processes, systems, and methods of operation (17 U.S.C. § 102(b)); (3) facts and functional classifications that merely track legal language; and (4) expression constrained by legal terminology and professional usage under merger and scènes-à-faire. What remains is thin protection under § 103(b) for any original phrasing or selection/arrangement. The court’s summary-judgment pattern—relief only where questions “very closely” tracked headnote prose—confirms that residue. *Mem. Op.* at 14.

The record places any copying at a nonpublic, intermediate step in a development pipeline. *Mem. Op.* at 18. Tokenization, lemmatization, and featurization (turning text into numbers) generate features and train parameters; the source text and Q–A files are then discarded. These are pipeline-bound copies aimed at function rather than expression. *ROSS SJ Opening Brief.*, D.I. 698 (D. Del. Oct. 1, 2024), at 23–26 (esp. 26).

On this record, users receive judicial opinions, not headnote prose or the Key Number taxonomy. Accordingly, any arguable appropriation lies, if at all, in the intermediate step; there is no output-level appropriation of protected headnote expression.

II. FACTOR ONE: NONPUBLIC, FUNCTIONAL TRAINING CAN BE JUSTIFIED IN THIS CASE

A. The Statute And Case Law Show That Intermediate, Nonpublic Copying To Reach Functionality Can Satisfy Factor One Even In Commercial Settings

Section 107 requires courts to assess “the purpose and character of the use” separately from “the effect of the use upon the market.” 17 U.S.C. § 107. *Warhol* articulates a practical presumption: when a secondary use shares the same or highly similar purpose as the original and is commercial, factor one will generally weigh against fair use absent some other justification for copying. 598 U.S. 508, 532–33 (2023) (paraphrased). That presumption does not end the inquiry, because *Campbell* makes clear that commerciality is a component of factor one, not a dispositive bar. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 584–85 (1994). This tracks the statutory structure, as the categorical separation of purpose from market effects means that market competition alone cannot decide factor one.

One “other justification” *Warhol* anticipates is supplied by the line of cases that permits nonpublic, intermediate copying to reach unprotected functionality:

Sega and *Sony/Connectix* allowed interim copying to access interfaces needed for interoperability; *Google v. Oracle* held that reusing functional declarations to allow users to carry over their skills can be fair use; and *Authors Guild v. Google, Inc.* and *HathiTrust* approved whole-work copying where the function was search, indexing, or accessibility and no protected expression was publicly communicated. *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1526 (9th Cir. 1992); *Sony Computer Entm't, Inc. v. Connectix Corp.*, 203 F.3d 596, 600–01 (9th Cir. 2000); *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 40 (2021); *Authors Guild v. Google, Inc.*, 804 F.3d 202, 207 (2d Cir. 2015) (hereinafter *Google Books*); *Authors Guild v. HathiTrust*, 755 F.3d 87, 97–98 (2d Cir. 2014).

The question may fairly be asked at this point: what logical structure in copyright requires this particular “other justification”? By examining the logic the Court used in *Warhol* to differentiate its “purpose or character of the use,” the requirement becomes clear. In *Warhol*, the Orange Prince screenprint that memorialized Prince on the cover of *Conde Nast* served both an identical purpose and commercial market—magazine cover licensing—as Goldsmith’s original photograph. But *Warhol* could not grant Goldsmith a monopoly over depictions of Prince; it merely protected her market for licensing her *expression* for covers. To claim otherwise—that the copyright covered the functional right to depict Prince—would give Goldsmith a potential claim not just against AWF, but against any

individual seeking to license any portrait of Prince to a magazine. Such a claim is antithetical to copyright law. Therefore, the law requires there be some “other justification” in cases concerning the copying of functional elements, which is exactly what *Google*, *Sega*, and *Sony* provide.

Seemingly contrary cases also track this logic upon closer inspection. In *TVEyes*, the Second Circuit rejected fair use because the service made available virtually all of the plaintiff’s audiovisual content to paying users; however, the court simultaneously indicated that the text-searchable database itself was outside the holding. *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 180–81 (2d Cir. 2018) (slip op. at 1 noting search not challenged). Even there, the court found the use transformative, ruling instead on factors three and four. *Id.* The Third Circuit’s *Video Pipeline* is of the same cloth: it found no transformation on the grounds that the defendant publicly streamed movie clips to the same audience and licensing market as the rightsholder’s trailers. *Video Pipeline, Inc. v. Buena Vista Home Entm’t, Inc.*, 342 F.3d 191, 199–200, 202–03 (3d Cir. 2003).

Taking the cases above, it appears courts apply the following two-part standard where the defendant’s copying serves a functional, intermediate purpose:

1. **Intermediate, purpose-distinct use.** The copying is part of a nonpublic process aimed at accessing ideas, facts, methods of operation, or interfaces

to build or improve a tool, rather than communicating the plaintiff's protected expression to the plaintiff's audience.

2. **No more than reasonably necessary.** The scope and manner of copying are tailored to that functional end: the defendant only extracts or implements the unprotected elements for the intermediate purpose.

“Reasonably necessary” has concrete content in the precedents. In *Sega* and *Connectix*, temporary copying of entire code images was allowed because there was no practical way to discover and implement unprotected interfaces without it, and the final products did not contain or reveal the copyrighted code. 977 F.2d at 1524–26; 203 F.3d at 607. In *Google v. Oracle*, reusing declarations—functional “methods of operation”—was fair because those declarations were needed so programmers could reuse their knowledge; Google wrote new implementation code rather than appropriating expressive implementation. 593 U.S. at 40. In *Google Books* and *HathiTrust*, whole-book scanning was fair where necessary to enable search or accessibility and protected text was not publicly exposed beyond non-substituting snippets or accessibility outputs. 804 F.3d at 207; 755 F.3d at 97–101. Across these decisions, necessity turns on whether the copying targets functionality and whether the taking is proportionate to the functional aim. As a brief side note, it is worth countering any confusion between prong two and factor three. While the two may sound similar, factor three analyzes the copying “in

relation to the copyrighted work as a whole,” whereas prong two focuses on analyzing the extent of the copying *in relation to the purpose of the new work*.

For a demonstration of this test, compare two hypothetical hiking apps. The first parses hiking-guide PDFs to extract GPS traces and display only the hiking trails to the end consumer. The second additionally republishes the guides’ narrative descriptions and ranking systems to users. The first passes the test (see *Perris v. Hexamer*, 99 U.S. 674 (1879) (map-maker has no exclusive copyright in map symbols)); the second does not. By parallel, copyright in headnotes protects editorial prose, not the functional mapping between legal ideas and authoritative passages.

In sum, where copying is a nonpublic, intermediate step to learn that mapping and the public output does not communicate headnote prose, *Warhol*’s presumption is overcome by the established “other justification.”

B. The Record Shows Nonpublic, Necessary Training That Extracts Functional Signals And Reveals No Headnote Prose, So Factor One Favors Fair Use

The record, as the court summarized it, identifies the “use at issue” as nonpublic model training within a development pipeline. The copying occurred at an intermediate step to fit a system that learns statistical relationships among legal concepts and authoritative passages. That pipeline relied on labeled

question-and-answer materials to supply the signal necessary for a learning-to-rank model. The text of those pairs was transformed into numerical features, the model's parameters were estimated from those features, and the source textual materials were not retained in the shipped product. The system's outputs to users are judicial opinions rather than headnote text. These facts, taken together, align with the purpose-and-character inquiry: the copying was directed at functionality in a nonpublic process, and the output does not communicate the plaintiff's expression to the public.

Against that backdrop, prong one is satisfied. The copying served an intermediate, purpose-distinct end: extracting functional relationships that allow a search tool to rank opinions responsive to user queries. The function is different in character from licensing headnote prose to legal researchers. The relevant audience for the challenged copying is not the public readership of headnotes, but a training process that tunes numerical parameters. The public-facing audience sees judicial opinions—unprotected law—rather than editorial prose. This is the same structural distinction that supported fair use in the search cases: *Google Books* emphasized that creating a search index serves a purpose different from reading the books, and the public interface did not disclose the protected text. 804 F.3d at 214–18. Likewise, *Sega* treated reverse engineering to access unprotected interfaces as functionally distinct from expressive copying, precisely because the public-facing

use did not communicate the plaintiff's code. 977 F.2d at 1524–26. Under *Warhol*, similarity of commercial field does not collapse the analysis where the *use at issue* is intermediate copying to reach functionality and the outputs do not transmit the plaintiff's expressive content.

Prong two is also satisfied. The steps described in the record track what is reasonably necessary for a learning-to-rank system to generalize from examples. Such a system requires labeled pairs linking queries to graded candidate answers in order to estimate parameters that will predict the relevance of new opinion passages. To fit those parameters, the system must parse the labeled text, transform it into numerical features, and run an optimization procedure that adjusts model weights to capture statistical relationships among terms, concepts, citations, and relevance labels. Interim access to the textual pairs is necessary to perform those steps. Once training is complete, those textual materials are no longer required for the system to operate, and the record indicates they are not retained in the released product. That tailoring—the minimal interim access required to extract functional signals, followed by discard of expressive text and outputs that do not reveal the plaintiff's prose—matches the necessity pattern approved in *Connectix*, *Google v. Oracle*, *Google Books*, and *HathiTrust*. 203 F.3d at 602–08; 141 S. Ct. at 1203–09; 804 F.3d at 217–25; 755 F.3d at 97–101.

The same facts that show necessity also confirm that the use remained within functional bounds. The public result of a user’s query is a ranked set of judicial opinions. That design is inconsistent with any purpose to communicate headnote prose and consistent with a training aim of learning functional mappings between questions and authorities. The training inputs were sought for their labels and relevance assessments—the signals the model needs—not for the editorial turns of phrase. The court’s recognition that the copying occurred “at an intermediate step” completes the picture: the activity is nonpublic, preparatory, and targeted at extracting unprotected relationships that allow a tool to function, not at substituting for licensed headnote expression.

With these record facts, *Warhol*’s presumption is overcome by the recognized justification for intermediate copying. The parties may compete within the legal research market, but under *Campbell* competition does not resolve factor one. 510 U.S. at 584–85. The statute reserves market effects for factor four, which is where courts assess whether a public-facing product usurps licensing markets for the plaintiff’s expressive work. That separation is not merely formalistic; it protects research, open projects, and smaller entrants who rely on predictable rules governing intermediate uses, while preserving a rigorous market analysis where it belongs.

C. Medium Labels Are Irrelevant; What Matters Is The Necessity Of Intermediate Copying To Reach Functionality

Characterizing this dispute as “not a code case” does not answer factor one. Section 107 is medium-agnostic. It asks how the defendant used the plaintiff’s work—whether the copying accessed functionality in a nonpublic, intermediate process and whether protected expression was communicated to the public—not whether the plaintiff’s work happens to be prose, code, or a hybrid system. New technologies often combine traits of prior media, and courts examine the retained characteristics that matter for the statute’s test rather than the label. *Sony Corp. of Am. v. Universal City Studios, Inc.* focused on the functional “time-shifting” characteristic of home video and the fact that a new copy was not being created for a new audience, not merely on “television” as a category. 464 U.S. 417, 466–67 (1984). The same approach applies here: model training on legal text straddles elements of prose and interface design, and the operative question is whether the copying was an intermediate means to reach unprotected functional relationships without conveying protected expression downstream.

Necessity is shown by the training design itself. A learning-to-rank system cannot be trained without labeled examples linking queries to graded answers. Those labeled pairs are the interface to functionality: they encode the relationships the system must learn to predict relevance. To extract those relationships, the

model must perform feature extraction and parameter estimation on the textual pairs. Interim access to the text or similar is required, just as *Sega* and *Connectix* needed to temporarily copy code to discover interfaces. 977 F.2d at 1526; 203 F.3d at 607. After training, retaining or exposing the source text is unnecessary, and the record does not indicate it occurred. *ROSS Opening Br.* pp. 23–26. That sequence—interim access to reach unprotected functionality, followed by discard and output constraints—tracks the necessity analysis in *Google v. Oracle*, where reusing declarations was required to allow users to carry over their skills, but wholesale copying of expressive implementation was not. 593 U.S. at 29–31. It also tracks *Google Books* and *HathiTrust*, where the scope of copying was justified by the function (search, indexing, accessibility) and was paired with strict limits on public exposure of expressive text. 804 F.3d at 217–19; 755 F.3d at 97–101.

Treating the medium label as dispositive would collapse the statute’s structure by importing market concerns into factor one simply because the parties sell research tools. *Warhol* instead directs courts to evaluate the *use at issue* and to ask whether, when purposes appear similar, there is “other justification for copying.” Here, the use is nonpublic, intermediate training to reach functional relationships that power search and issue-spotting; the outputs are judicial opinions; and the record indicates no mechanism by which headnote prose is stored in or displayed by the shipped system. On those facts, the established justification

for intermediate copying applies regardless of whether the plaintiff’s input materials resemble code, prose, or both. Whether the resulting product affects any licensing market for headnote expression is a separate factor-four question.

Preserving that allocation respects the text of § 107, maintains coherence with *Sega*, *Connectix*, *Google v. Oracle*, *Google Books*, and *HathiTrust*, and supplies predictable guidance: intermediate, nonpublic copying to reach functionality, limited to what is reasonably necessary favors fair use under factor one; any residual market-effect arguments are examined under factor four.

III. FACTOR FOUR SAYS TO COUNT HARM ONLY TO MARKETS FOR PROTECTED EXPRESSION; NONE IS SHOWN HERE

A. The Relevant Market Is Licensing Of Headnote Expression—Not The Platform-level Market For Legal Research

17 U.S.C. § 107(4), read with 17 U.S.C. §§ 102(b), 103(b), counts harm to markets that traffic in the work’s protected expression and traditional or likely derivative licenses for that expression, not to ideas, facts, methods of operation, or access to public-domain law (*Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 348–51 (1991); *Baker v. Selden*, 101 U.S. 99, 103–04 (1880)). Properly defined here, the market is licensing or replacement of headnote text (and any established republication licenses), not the broad, functional market for “legal-research platforms.” Defining the market at the level of methods would tax

research itself, not expression, and would slow domestic development of lawful AI systems.

Courts have already addressed closely analogous contexts. Search-oriented systems that provide non-expressive outputs (e.g., indices, snippets, pointers) do not usurp markets for the underlying expressive texts. *Google Books*, 804 F.3d at 223–25; *HathiTrust*, 755 F.3d at 99–101; *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1168–69 (9th Cir. 2007) (search and thumbnails facilitate location; no cognizable harm to markets for full-sized images). Those cases cabin factor-four analysis to expressive licensing markets actually implicated by the defendant’s outputs.

Defining the market as “legal research” is legally incorrect on four independent grounds. First, it protects methods—search, ranking, issue-spotting—that § 102(b) excludes. *Feist*, 499 U.S. at 348–51; *Baker*, 101 U.S. at 103–04. Second, it conflates a product market (platform rivalry) with a permissions market (licenses to republish expression), which factor four does not do. Third, it collapses Factors One and Four by re-labeling functional competition as “market harm.” Fourth, it violates *Bill Graham*’s bar on circular derivative markets by defining a market by the challenged use itself (here, the training licenses). *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 614–15 (2d Cir. 2006).

Warhol reinforces this framing: factor analysis respects traditional, reasonable, or likely markets actually served by the original, not bespoke markets reverse-engineered from the defendant's technology. 598 U.S. at 532–34. The record's output facts control: users receive judicial opinions and citations, not headnote prose. That narrows any cognizable market to headnote expression and its licensed republication, if any.

B. Courts Find Harm Through Substitution, Usurpation Of Expressive Licensing Markets, Or Erosion Of Price or Exclusivity; None Is Present Here

There is no presumption of market harm. The copyright owner bears the burden to show a meaningful likelihood that the defendant's outputs will substitute for licensed publication of the plaintiff's expressive text, with evidence—not speculation. *Campbell*, 510 U.S. 569, 593–94.

Courts credit harm when (i) the secondary use substitutes for the plaintiff's expressive work or licensed derivatives (*Campbell*, 510 U.S. 569, 590–91 (1994); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 821–22 (9th Cir. 2003); see also *Harper & Row, Publ'rs, Inc. v. Nation Enters.*, 471 U.S. 539, 566–69 (1985)); (ii) it usurps an established or likely permissions market for that expression (*Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930–31 (2d Cir. 1995); *Princeton Univ. Press v. Mich. Document Servs., Inc.*, 99 F.3d 1381, 1389–91 (6th Cir. 1996) (en banc); *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 178–81 (2d Cir. 2018)); or

(iii) it erodes price or exclusivity of such expressive licenses (*Harper & Row*, 471 U.S. at 566–69; *Salinger v. Random House, Inc.*, 811 F.2d 90, 99–100 (2d Cir. 1987)). Only traditional, reasonable, or likely to develop markets qualify, speculative or circular “derivative markets” do not (*Castle Rock Entm’t, Inc. v. Carol Publ’g Grp., Inc.*, 150 F.3d 132, 145–46 (2d Cir. 1998); *Bill Graham*, 448 F.3d 605, 614–15 (2d Cir. 2006)).

Applied here, none of the three avenues is satisfied:

No substitution. Outputs contain no headnote text and thus cannot have competed with or displaced licensed publication of headnote prose. Users received opinions and citations which serve location and analysis, not expressive replacement—precisely the kind of non-substitutive functionality courts have approved. *Google Books*, 804 F.3d at 223–25; *HathiTrust*, 755 F.3d at 100–02; *Perfect 10*, 508 F.3d at 1168–69. At scale, the effect is more competition in method, not less demand for expressive headnote outputs. *Campbell*, 510 U.S. at 590 (no “meaningful likelihood” of substitution).

No usurpation of a functioning expressive license. Plaintiff identifies no established license by which it authorizes third-party public republication of headnote prose (as opposed to product subscriptions to its platform). In fact, ROSS cites Thomson Reuters’ discovery responses (RFA 72; Interrog. 21) stating it has never taken steps to sell Westlaw content as AI training data; see *ROSS Opening*

Br. Statement of Facts § C, pp. 10–11, 17–18. *Texaco* and *Princeton* anchored harm in recognized permissions markets to copy and distribute expressive texts; the asserted market here is platform sales or an abstract training license, not a license to publish headnote expression to the public. 60 F.3d at 930-91, 99 F.3d at 1389-91. *Video Pipeline* is illustrative of the harm that counts here: public streaming of clips usurped the rightsholder’s established trailer-licensing market and the value the rightsholder was deriving from that market, tipping factor four. 342 F.3d at 201–05. By contrast, ROSS’s opinion-only outputs do not displace any license to publish headnote prose.

No erosion of price/exclusivity. Where users receive only opinions/citations and not headnote text, there is no basis to claim diminished willingness to pay for, or impaired exclusivity of, any license to republish headnote expression. The relevant price/exclusivity attaches to expressive headnote licensing; absent public output of headnote text, there is nothing to erode. *Harper & Row*, 471 U.S. at 566–69; *Salinger*, 811 F.2d at 99–100.

By contrast, if—contrary to the present record—ROSS’s search tool additionally supplied an auto-generated taxonomy that substituted for West’s headnotes, factor four would cut in favor of Thomson Reuters, even with some language variation. West’s headnotes and the Key Number taxonomy are expressive choices in selection, phrasing, and arrangement. A ROSS product that

relied on the heart of that work to publish a competing set of labeled summaries or doctrinal bins to end-users would target the same market as West’s headnotes; it would be a classic case of expressive-output substitution or usurpation of a traditional permissions market if one were established.

C. Non-cognizable Theories Including “Training-license” Markets And “Market Dilution” Should Be Rejected

A bare “license to be trained on” regulates internal method, not public exploitation of expression. Because 17 U.S.C. § 102(b) excludes ideas, procedures, processes, systems, and methods of operation from copyright protection, a training-only license that attempts to define the relevant market effectively seeks to convert an unprotected method into a copyrighted market. It therefore defines the market by the challenged practice rather than by any established license to publish protected text. That is circular and non-cognizable under factor four. *Bill Graham*, 448 F.3d at 614–15. Plaintiffs cannot conjure a new training license market post-hoc and claim harm to it.

The plaintiff may respond that its headnote system is “expressive and valuable.” Value alone is irrelevant absent expressive substitution or usurpation of an established expressive license. The law protects headnote expression, not a monopoly over research methods or over access to judicial opinions. *Feist*, 499 U.S. at 348–51; *Banks*, 128 U.S. at 253–54.

So too with “market dilution” theories. The *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417-VC (N.D. Cal. June 25, 2025) dicta suggesting that LLMs might “flood the market” with works “in a style” even without copying text protects ideas, styles, and tropes, not expression. That theory is circular (it defines a market by the plaintiff’s complaint), collapses Factor One back into Factor Four (counting idea-level competition as “indirect substitution”), misreads substitution doctrine (which asks whether the defendant’s output replaces the plaintiff’s expression or an established expressive license), and lacks a limiting principle. Copyright does not insure against more works in a genre; it protects against expressive displacement. On a record of opinion-only outputs and nonpublic intermediate training, *Kadrey*’s dilution concept has no purchase.

Finally, to the extent the plaintiff gestures at “likely to be developed” markets, the qualifier is not a blank check. Courts look for markets that are reasonable and proximate to the original’s exploitation—coursepack licenses for articles, image licensing for photographs, magazine licensing for portraits—not abstract permissions divorced from public expressive publication. *Texaco*, 60 F.3d at 930–31; *Princeton Univ. Press*, 99 F.3d at 1389–91; *Warhol*, 598 U.S. at 532–34. A nonpublic “training” permission aimed at internal model improvement is not a license to publish headnote prose; counting it would collapse factor four into a tax on method.

D. Policy Note

Treating nonpublic “training” as a cognizable market would misalign copyright with its constitutional purpose of promoting progress by protecting expression while leaving ideas, facts, and methods free for all. It would balkanize datasets, raise fixed compliance costs, and entrench incumbents at the expense of research labs, nonprofits, and startups. Consider a two-engineer startup that wants to fine-tune a model on the scientific and news literature relevant to its product. It would have to chase dozens of bespoke “training” permissions with incompatible terms and fees, build compliance tooling it cannot afford, and likely abandon the release, while an incumbent with a licensing department proceeds. That is a toll on method, not expression.

The fragmentation problem is particularly acute in law. The public needs reliable access to judicial opinions; editors add value by writing headnotes, but that value does not extend to controlling how others search, rank, or cluster public-domain texts. Counting internal training as a factor-four “market” would make the means of reading law licensable, not the expression of headnotes. That inversion would reduce competition in research features, slow the diffusion of quality tools beyond elite institutions, and increase costs for courts, clinics, and small firms that rely on affordable search.

The alternative is administrable and protective: keep factor four tethered to the public exploitation of expression, and police outputs. If a system emits headnote prose, that is classic substitution and can be enjoined or licensed. If it does not—if it outputs opinions/citations and analytic signals—then the competition is in method, not in the market for expressive headnote publication. This narrow rule protects nonpublic, necessary training while policing public outputs and advances the broader national interest in AI leadership without sacrificing authors’ expressive rights.

On this record, the findings that (i) ROSS’s outputs exclude headnote text and (ii) the search market fosters competition in methods rather than expressive content weigh against any cognizable harm to expressive or established licensing markets and instead point to functional effects that factor four does not address.

CONCLUSION

For the reasons set out in the two-factor discussion above—and given the district court’s prior acceptance that factors two and three tilt toward fair use—the Court should adopt the proposed factor-one standard (nonpublic, purpose-distinct intermediate use; and copying no more than reasonably necessary) and hold that ROSS’s use is fair use as a matter of law under § 107.

Alternatively, if the Court concludes a genuine dispute remains under factor four, the case should be sent to the jury solely to determine the extent of any

market impact on licensing for the headnotes’ expressive content, not platform-level competition, methods of operation, or bare “training-license” theories.

Accordingly, the Court should reverse and render judgment for ROSS on fair use; or, barring that, vacate and remand for a jury trial limited to the factor four question described above.

September 29, 2025

Respectfully submitted,

/s/ Tim Hwang

Tim Hwang

FOUNDATION FOR
AMERICAN INNOVATION

2443 Fillmore St #380-3386

San Francisco, CA, 94115

Telephone: (973)-960-4955

tim.hwang@thefai.org

ADDENDUM: CONSENT OF THE PARTIES (FRAP 29(a)(2))

Amici curiae state that all parties have consented to the filing of this brief pursuant to Federal Rule of Appellate Procedure 29(a)(2). On July, 21, 2025, the parties agreed to provide blanket consent to amicus filings on appeal subject to seven (7) days' advance notice. On September 22, 2025, at 11:42 a.m., amici provided written notice of their intent to file.

This Corrected Amicus Curiae Brief is identical in all substantive respects to the brief previously filed on September 29, 2025 (ECF No. 50) except that the attached "Addendum: Consent of the Parties (FRAP 29(a)(2))" has been appended. No other changes have been made to the brief's text, citations, arguments, or formatting.

CERTIFICATE OF COMPLIANCE

I hereby certify the following:

1. I am a member in good standing of the bar of the United States Court of Appeals for the Third Circuit. 3d Cir. L.A.R. 28.3(d).
2. This brief complies with the type-volume limitation of F.R.A.P. 29(a)(5) and 32(a)(7)(B) because this brief contains 5,302 words, excluding the parts of the brief exempted by F.R.A.P. 32(f) and 3d Cir. L.A.R. 29.1(b).
3. I served a copy of this brief on all counsel for the parties electronically through this Court's docketing system.
4. The text of the electronic brief is identical to the text in the paper copies.
3d Cir. L.A.R. 31.1(c).
5. The VirusTotal virus protection program, has been run on this file and no virus was detected. 3d Cir. L.A.R. 31.1(c).
6. This brief complies with the typeface and type-style requirements of F.R.A.P. 32(a)(5) and (a)(6) because it has been prepared in a proportionally spaced typeface using Microsoft Word in fourteen-point Times New Roman font.

September 29, 2025

Respectfully submitted,

/s/ Tim Hwang

Tim Hwang

Foundation for American Innovation

2443 Fillmore St #380-3386

San Francisco, CA, 94115

Telephone: (973)-960-4955

tim.hwang@thefai.org

CERTIFICATE OF SERVICE

I certify that on September 30, 2025, the foregoing motion was filed using the Court's CM/ECF system. All participants in the case are registered CM/ECF users and will be served electronically via that system.

September 30, 2025

Respectfully submitted,

/s/ Tim Hwang

Tim Hwang

FOUNDATION FOR

AMERICAN INNOVATION

2443 Fillmore St #380-3386

San Francisco, CA, 94115

Telephone: (973)-960-4955

tim.hwang@thefai.org