No. 25-2153

IN THE UNITED STATES COURT OF APPEALS
FOR THE THIRD CIRCUIT

THOMSON REUTERS ENTERPRISE CENTRE GMBH;
WEST PUBLISHING CORP.,
Plaintiffs-Appellees,

v.

ROSS INTELLIGENCE INC.,
Defendant-Appellant.

On Appeal from the United States District Court
for the District of Delaware
No. 1:20-cv-00613, Judge Stephanos Bibas (sitting by designation)

BRIEF OF AMICI CURIAE ABRAHAM KANG AND KUNAL PATEL IN SUPPORT OF APPELLANT ROSS INTELLIGENCE INC.'S PETITION FOR PANEL REHEARING OR REHEARING EN BANC

Abraham Kang
Counsel of Record for Amici Curiae
15732 Los Gatos Blvd #5002
Los Gatos, CA 95032
(650) 600-0178
abejkang@gmail.com

Kunal Patel
Amicus Curiae

DISCLOSURE STATEMENT

Amici Abraham Kang and Kunal Patel are individuals. Neither is a corporation, so no corporate disclosure statement is required. Fed. R. App. P. 26.1(a), 29(a)(4)(A).

TABLE OF CONTENTS

Page

STATEMENT OF IDENTITY, INTEREST, AND AUTHORITY TO FILE 1

INTRODUCTION AND SUMMARY OF ARGUMENT 2

ARGUMENT 3

I. A Headnote Framed a Question; the Grades Did the Teaching. 3

A. The record describes a standard graded-relevance dataset. 3

B. Training draws on the point of law, not the editor's words. 4

C. The record divides the work between two companies. 5

II. Sega's "Necessity" Was Practical and Went to the Work's Nature. 6

A. The panel held that ease never justifies copying. 6

B. The distinguished cases did not hold that. 7

C. Both courts weighed necessity under factor two. 7

D. Building labeled data is the same kind of practical choice. 8

III. What Reached the Public, and What the Market Lost. 9

A. ROSS's output could not reproduce the headnotes at issue. 9

B. ROSS's users received a search over public opinions, not headnotes. 9

C. The counted license was not for sale to competitors. 10

IV. Footnote 7 Limits the Facts, Not the Tests. 11

A. The footnote distinguishes outputs; the tests judge training. 11

B. Modern AI systems, including chatbots, learn from paired examples.....12

C. The panel can confine its tests; otherwise the full Court must.....13

CONCLUSION.....14

CERTIFICATE OF BAR MEMBERSHIP 15

CERTIFICATE OF COMPLIANCE 16

TABLE OF AUTHORITIES

Page(s)

Cases

Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508 (2023) 5, 6

Fox News Network, LLC v. TVEyes, Inc., 883 F.3d 169 (2d Cir. 2018) 9

Sega Enterprises Ltd. v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992) 2, 6, 7, 11

Sony Computer Entertainment, Inc. v. Connectix Corp., 203 F.3d 596 (9th Cir. 2000).....2, 7, 8, 11

Thomson Reuters Enter. Centre GmbH v. ROSS Intel. Inc., No. 25-2153 (3d Cir. Sept. 29, 2026) ..... passim

Video Pipeline, Inc. v. Buena Vista Home Ent., Inc., 342 F.3d 191 (3d Cir. 2003).....10

Statutes

17 U.S.C. § 102(b) 10

Rules

Fed. R. App. P. 29(a)(4)(E) 2

Fed. R. App. P. 29(b)(2)..... 1

3d Cir. I.O.P. 9.1 13

3d Cir. I.O.P. 9.5.9 13

Other Authorities

Long Ouyang et al., *Training Language Models to Follow Instructions with Human Feedback* (Mar. 4, 2022), https://arxiv.org/abs/2203.02155 12

STATEMENT OF IDENTITY, INTEREST, AND AUTHORITY TO FILE

Amicus Abraham Kang is an attorney admitted to the bar of this Court and a computer scientist whose research is in natural language processing and machine learning. Amicus Kunal Patel is an engineer who designs and evaluates machine-learning systems trained on labeled examples. Amici filed a merits-stage brief supporting ROSS. ECF No. 72.

The panel's tests govern how training data may be built and used, which is amici's daily work. Three things concern amici most: training data usually comes from many owners, and clearing every one is often impossible; the market test would let whoever holds a field's data refuse a license, train its own models, and block a newcomer's training by counting it as the harm; and training is the same step whether a model ranks passages or writes text. Where this brief explains how such systems are generally built, it offers amici's professional experience, not findings about ROSS.

Amici seek leave to file by the accompanying motion. Fed. R. App. P. 29(b)(2). No party's counsel authored this brief in whole or in part. No party or party's counsel contributed money intended to fund preparing or

submitting it, and no person other than amici contributed such money. Fed. R. App. P. 29(a)(4)(E).

INTRODUCTION AND SUMMARY OF ARGUMENT

The panel held that ROSS's use of Westlaw headnotes to train an AI model was not fair use. *Thomson Reuters Enter. Centre GmbH v. ROSS Intel. Inc.*, No. 25-2153, slip op. 27 (3d Cir. Sept. 29, 2026). ROSS asks for rehearing en banc or, in the alternative, panel rehearing. Appellant's Pet. for Panel Reh'g or Reh'g En Banc 19, ECF No. 225 ("Pet."). Amici support either course. They do not repeat the petition's doctrine. Pet. 7–16. They offer an engineer's account, from the record, of what a training example does and why tests stated in general terms reach systems the panel did not set out to decide. *See* slip op. 2, 17 n.7.

First, a headnote helped a person write a question; the model learned from grades attached to public court passages (Part I).

Second, *Warhol* asks whether copying is "reasonably necessary," and the necessity in *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), and *Sony Computer Entertainment, Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000), was practical and went to the nature

of the work. Neither case held that "ease is not a justification for copying." Slip op. 21 (Part II).

Third, ROSS's output could not reproduce the 2,243 headnotes at issue, and the license the panel counted is one Thomson Reuters would not sell to a competitor (Part III).

Fourth, footnote 7 limits the facts, not the tests. The full Court should settle the tests and return the case to the panel to apply them (Part IV).

None of this limits an owner's remedies against outputs that reproduce protected expression. *See* Pet. 18.

ARGUMENT

I. A Headnote Framed a Question; the Grades Did the Teaching.

A. The record describes a standard graded-relevance dataset.

ROSS "enlisted a company called LegalEase Solutions," which "wrote approximately 25,000 legal memos, each of which presented a legal question and four to six judicial opinion passages in response." Slip op. 5. Memo writers "used thousands of Westlaw headnotes to create the memos' legal questions." *Id.* at 5–6. Each memo had to include "at least one great, good, topical, and irrelevant judicial opinion passage, labeled

as such." *Id.* at 6. The memos were then "converted into a machine-readable format and used to train the ROSS AI program to identify patterns regarding which judicial opinion passages respond well to legal questions." *Id.*

In amici's experience, that is a standard graded-relevance dataset: a question, several candidate passages, and a grade for each. Such data typically trains a ranking model. The model reads a question and one passage and outputs a single number: how well the passage answers the question. In training, those numbers are compared with the human grades, and the model's internal settings, called parameters, are adjusted so that "great" passages score higher than "irrelevant" ones. The trained model keeps the adjusted parameters, not a file of the questions.

B. Training draws on the point of law, not the editor's words.

A headnote is written for a human reader. Each must "make sense to a reader who doesn't have the opinion." Slip op. 4. In a training example, the question does a different job. It is a probe. The model is never told whether the question states the law correctly, only which public passages respond to it, and how well.

What the question carries is the legal issue, and the issue is the court's. Two publishers headnoting the same opinion state the same point of law in different words. *See* Lexis Br. 13–16, ECF No. 100. In amici's experience, the same issue asked in other words, with the same passages graded alike, teaches a ranking model the same thing. The editor's wording is what training needs least.

C. The record divides the work between two companies.

Stage one belonged to LegalEase and its subcontractor, which used headnotes to frame questions and paired each with graded public passages. Slip op. 5–6. Stage two belonged to ROSS, which trained the model. *Id.* at 6. The panel judged the first stage's copying by the second stage's product, not its own: ROSS "used the headnotes to train an AI program for the benefit of its legal-research platform," so its use "shares the same ultimate purpose as Thomson Reuters's use." *Id.* at 16–17.

A graded dataset is reusable: the same memos could train a tutoring tool or serve as a research benchmark. Under the panel's test, the same copying would fare differently at factor one depending on what the memos later trained. The Supreme Court asks instead about "the specific 'use' of a copyrighted work that is alleged to be 'an infringement.'"

Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 533 (2023). Here, that use built a training set: questions paired with graded answers, not a search engine.

ROSS does not contest that LegalEase's acts are attributable to it. Slip op. 6 n.2. Attribution decides who answers for a copy, not what the copy was used for.

II. Sega's "Necessity" Was Practical and Went to the Work's Nature.

A. The panel held that ease never justifies copying.

The panel distinguished *Sega* and *Connectix* because "in each case, copying was necessary to access the unprotected functional aspects of computer code." Slip op. 20. ROSS, by contrast, copied because the headnotes "offered an 'easy' way to create its training memos," and "[u]nlike necessity, ease is not a justification for copying." *Id.* at 20–21. The same rule decided factor three: copying was "not necessary" because the opinions "were freely available." *Id.* at 23. *Warhol* asks instead whether copying is "reasonably necessary to achieve the user's new purpose," a question of degree, not whether another route existed. 598 U.S. at 532.

B. The distinguished cases did not hold that.

Sega agreed that "[t]rained programmers can disassemble object code by hand." *Sega*, 977 F.2d at 1525. It found the copying justified anyway, because "even a trained programmer cannot possibly remember the millions of zeros and ones that make up a program." *Id.*

In *Connectix*, an engineer admitted that copying "was not the only way to fix the bug, just the fastest way to do so." *Connectix*, 203 F.3d at 604–05 n.8. Finding fair use anyway, the court declined a rule that "would require that a software engineer, faced with two engineering solutions that each require intermediate copying of protected and unprotected material, often follow the least efficient solution." *Id.* at 605.

Both routes in *Connectix* involved some copying; the panel's alternative here would have copied no headnotes. But the panel stated the general rule above, one *Connectix* rightly declined to adopt.

C. Both courts weighed necessity under factor two.

Connectix discussed necessity under factor two, the "[n]ature of the copyrighted work." *Id.* at 603. Because Sony's BIOS "contains unprotected aspects that cannot be examined without copying," the court accorded it "a 'lower degree of protection than more traditional literary

works,'" and so required that the copying be "'necessary' to have been fair use." *Id.* (quoting *Sega*, 977 F.2d at 1526). The panel weighed necessity under factors one and three instead, though it held that factor two favors fair use because headnotes "must accurately convey the law and describe uncopyrightable judicial opinions." Slip op. 20–23. Like Sony's BIOS, then, headnotes lie "at a distance from the core." *Connectix*, 203 F.3d at 603.

D. Building labeled data is the same kind of practical choice.

In amici's experience, labeled examples are often the most expensive input to a machine-learning system, and writing realistic questions is the slowest part of making them. Teams commonly seed questions from short existing descriptions of documents, such as titles, abstracts, or summaries, to produce tens of thousands of examples at a workable cost. ROSS's memos numbered "approximately 25,000." Slip op. 5. Each training run is an experiment; a working model often takes a dozen or more, many of them gathering or revising data, so the final dataset can vastly understate the data effort behind it. If "ease is not a justification for copying," *id.* at 21, the only safe way to build such data is the slowest one.

Amici ask only that the ease rule be withdrawn. They do not ask the Court to find that ROSS met any necessity standard.

III. What Reached the Public, and What the Market Lost.

A. ROSS's output could not reproduce the headnotes at issue.

The 2,243 headnotes at issue are those whose memo questions were "so similar to the headnote text but so dissimilar to the opinion text." Slip op. 7 n.3. ROSS's system "would only return text passages from preexisting judicial opinions." *Id.* at 5; *see* A55. Its output, which is opinion text, therefore could not reproduce those headnotes.

On factor three, a decision the panel cited explains that "[t]he relevant consideration is the amount of copyrighted material made available to the public rather than the amount of material used by the copier." *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 179 (2d Cir. 2018); see slip op. 18. For the headnotes at issue, that amount was zero. See Pet. 12–14.

B. ROSS's users received a search over public opinions, not headnotes.

On Westlaw, a headnote sits above the opinion, links to the passage it reflects, and ties into the Key Number System that indexes opinions by

topic. Slip op. 3. ROSS offered none of that. It answered plain-language questions with passages from public opinions. *Id.* at 5. ROSS competed with Westlaw, *id.* at 19, but what its users received was not a headnote.

The panel found that ROSS "appropriated the headnotes' value for finding and analyzing judicial opinions." *Id.* at 25. Helping researchers find and analyze judicial opinions is a function that copyright does not itself protect. *See* 17 U.S.C. § 102(b). In *Video Pipeline*, by contrast, the harm ran through public display of the owner's own expression: retail websites paid to display clip previews made from Disney's films. *Video Pipeline, Inc. v. Buena Vista Home Ent., Inc.*, 342 F.3d 191, 202 (3d Cir. 2003).

C. The counted license was not for sale to competitors.

The panel held that ROSS "usurped Thomson Reuters's opportunity to enter that derivative market and license its headnotes for that purpose." Slip op. 26. But Thomson Reuters "does not let users use Westlaw to develop a competing platform," A3, and when ROSS asked for a license, refused one "because Ross was its competitor," A37. It trains its own AI search products on headnotes. Slip op. 4, 26.

On this record, the license the panel counted is one a competitor could not have bought. What ROSS's use cost Thomson Reuters was the ability to exclude a competitor, and copyright law "does not confer such a monopoly." *Connectix*, 203 F.3d at 607. A monopoly over a work's function is patent law's to grant, and only on its "more stringent standards." *Sega*, 977 F.2d at 1526.

IV. Footnote 7 Limits the Facts, Not the Tests.

A. The footnote distinguishes outputs; the tests judge training.

Footnote 7 distinguished the generative-AI cases because ROSS's platform "cannot generate original expression" and because ROSS trained its AI "for the purpose of creating a commercial substitute for Westlaw." Slip op. 17 n.7. But none of the panel's three tests depends on what a model outputs. They ask whether a training copy "shares the same ultimate purpose" as the owner's use, *id.* at 17; whether ease can justify copying, *id.* at 21; and whether copying "for use as training data without Thomson Reuters's authorization" took a licensing opportunity, *id.* at 26.

B. Modern AI systems, including chatbots, learn from paired examples.

The panel defined AI as a program "given data which contains examples of inputs paired with outputs" that learns "to recognize patterns regarding which outputs are responsive to which inputs." Slip op. 5 n.1. In amici's experience, that describes how large language models are trained at every stage. In pretraining, a stretch of text is the input and the word that follows is the output. In instruction tuning, a prompt is the input and a written model answer is the output. In preference training, people rank several candidate answers to the same prompt, and the model learns to favor the higher-ranked ones.

That last step has the same shape as ROSS's memos. One widely cited account reports: "In order to speed up comparison collection, we present labelers with anywhere between K = 4 and K = 9 responses to rank." Long Ouyang et al., *Training Language Models to Follow Instructions with Human Feedback* 8 (Mar. 4, 2022), https://arxiv.org/abs/2203.02155. ROSS's memos paired each question with "four to six" graded passages. Slip op. 5. One prompt, several candidates, and a human judgment about which is best: the panel's tests

would judge that step the same way whether the model ranks passages or writes text.

C. The full Court should settle the tests; the panel can apply them.

The panel called this "no more than an ordinary copyright case." Slip op. 2. Its tests, as written, reach further. A later panel applying them to a chatbot or a fraud filter will find nothing in them that turns on output, and in this Court "the holding of a panel in a precedential opinion is binding on subsequent panels." 3d Cir. I.O.P. 9.1.

The tests run through factors one, three, and four, so an amendment is unlikely to fix them all. And once they stand, "Court en banc consideration is required" to revise them. 3d Cir. I.O.P. 9.1. The full Court should take that step now. It can grant rehearing "as to one or more of the issues," 3d Cir. I.O.P. 9.5.9, settle the tests, and return the case to the panel to apply them. The tests should judge each training copy by its own use, building a dataset, not by the product the model later serves; weigh necessity by degree, as *Warhol* does; and count substitution for headnotes, not an unsold license.

CONCLUSION

The Court should grant rehearing en banc, settle the tests for copies made to train AI models, and return the case to the panel to apply them. Failing that, it should grant panel rehearing.

Dated: November 10, 2026

Respectfully submitted,

/s/ Abraham Kang
Abraham Kang
Counsel of Record for Amici Curiae
15732 Los Gatos Blvd #5002
Los Gatos, CA 95032
(650) 600-0178
abejkang@gmail.com

Kunal Patel
Amicus Curiae

CERTIFICATE OF BAR MEMBERSHIP

I certify that 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), 46.1(e).

Dated: November 10, 2026

Respectfully submitted,

/s/ Abraham Kang

Abraham Kang
Counsel of Record for Amici Curiae
15732 Los Gatos Blvd #5002
Los Gatos, CA 95032
(650) 600-0178
abejkang@gmail.com

CERTIFICATE OF COMPLIANCE

1. This brief complies with the type-volume limit of Federal Rule of Appellate Procedure 29(b)(4) because it contains 2,362 words, excluding the parts of the brief exempted by Federal Rule of Appellate Procedure 32(f): the cover page, disclosure statement, table of contents, table of authorities, signature block, and certificates. Headings and the statement required by Rule 29(a)(4)(E) are included in the count. The count was made with the word-count feature of Microsoft Word.

2. This brief complies with the typeface and type-style requirements of Federal Rule of Appellate Procedure 32(a)(5) and (6) and 3d Cir. L.A.R. 32.1(c) because it has been prepared in a proportionally spaced serif typeface, Century Schoolbook, 14-point, using Microsoft Word.

3. Under 3d Cir. L.A.R. 31.1(c), the text of the electronic brief is identical to the text of any paper copies filed with the Court.

4. Under 3d Cir. L.A.R. 31.1(c), a virus-detection program, Avast Security, version 16.5.0g, was run on the electronic file of this brief, and no virus was detected.

Dated: November 10, 2026

Respectfully submitted,

/s/ Abraham Kang

Abraham Kang
Counsel of Record for Amici Curiae
15732 Los Gatos Blvd #5002
Los Gatos, CA 95032
(650) 600-0178
abejkang@gmail.com