UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE OPENAI, INC., COPYRIGHT INFRINGEMENT LITIGATION This document relates to: No. 1:23-cv-08292 (SHS) (OTW) No. 1:23-cv-10211 (SHS) (OTW) No. 1:25-cv-03482 (SHS) (OTW)	Case No. 25-md-3143-SHS-OTW Hon. Sidney H. Stein **REDACTED – FILED PUBLICLY**

CLASS PLAINTIFFS' REPLY IN SUPPORT OF SUMMARY JUDGMENT

TABLE OF CONTENTS

INTRODUCTION **1**

FACTUAL BACKGROUND **3**

ARGUMENT **3**

I. DEFENDANTS' CLAIMED "ULTIMATE" PURPOSE DOES NOT CONTROL BECAUSE THE COURT MUST ANALYZE ALL THREE OF DEFENDANTS USES SEPARATELY. 3

A. Distinct Uses Require Distinct Treatment and Defendants' "Copy-By-Copy" Argument Rebuts a Straw-Man. 4

B. Defendants' Training Use Cannot Subsume Their Separate Piracy and Distribution Uses. 5

i. *Texaco* Bars Reliance on Defendants' Claimed Downstream "Training" Purpose to Vindicate Unpaid Downloading or Distributions. 5

ii. Defendants' Remaining Citations Confirm Subsequent Use Cannot Retroactively Justify Pirated Downloading or Distributions. 6

II. USE #1: OPENAI'S FREE DOWNLOADS OF MILLIONS OF BOOKS FROM NOTORIOUS PIRATE SITES WAS INFRINGEMENT AS A MATTER OF LAW. 8

A. Factor One: OpenAI's Purpose in Pirating Millions of Books Was Substitutive, Not Transformative. 8

B. Factor Four: OpenAI's Free Downloads Harmed Plaintiffs' Directly and Would Destroy Their Traditional Markets If "Unrestricted and Widespread." 10

III. USE #2: FAIR USE DOES NOT PROTECT OPENAI'S SUBSTITUTIVE, COMMERCIAL, AND UNNECESSARY USE OF PLAINTIFFS' WORK TO TRAIN LLMS. 11

A. Factor One: OpenAI Used Plaintiffs' Works to Optimize Its Models Towards Generating Competing Output for Cheap. 12

i. Factor One Favors Plaintiffs Because Defendants' Purpose Was Objectively Substitutive. 13

ii. Factor One Favors Plaintiffs Because Defendants Lack Justification for Their Use of Plaintiffs' Works. 15

iii. At Minimum, Factual Disputes Regarding the Nature of LLM Training Precludes Summary Judgment in Defendants' Favor. 17

B. Factor Four: Displacement Caused By The Flood of AI-Generated Works and Harms to the Already-Existing AI Training Licensing Market Weigh Strongly Against Fair Use For Training. 18

i. The Flood of AI-Generated Works Harms Plaintiffs. 18

ii. The Market to License Books for AI-Training Exists And Is Rapidly Expanding. 19

iii. The Court Must Consider Harms from Market Displacement As a Matter of Law. 20

iv. The Court Must Consider Harms from Lost Licensing Fees as a Matter of Law. 23

IV. USE #3: DEFENDANTS' UNPAID "HORSE-TRADING" DISTRIBUTIONS OF PLAINTIFFS' WORKS WERE NOT FAIR USE. 24

A. Distribution is a Separate Use From Initial Acquisition or Training And is, Independently, Not Fair. 24

B. OpenAI's Download of [redacted] Distributions to Contractors, and [redacted] Are Infringing Conduct At Issue In This Case. 25

V. MICROSOFT IS VICARIOUSLY LIABLE FOR OPENAI'S INFRINGEMENT. 26

CONCLUSION 26

TABLE OF AUTHORITIES

Page(s)

Cases

A&M Recs., Inc. v. Napster, Inc.,
239 F.3d 1004 (9th Cir. 2001) 6, 8

American Geophysical Union v. Texaco,
60 F.3d 913 (2d Cir. 1994)..... *passim*

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

Arica Inst., Inc. v. Palmer,
970 F.2d 1067 (2d Cir. 1992).....14, 15, 21, 22

Authors Guild, Inc. v. HathiTrust,
755 F.3d 87 (2d Cir. 2014).....15, 24

Authors Guild v. Google, Inc.,
804 F.3d 202 (2d Cir. 2015)..... *passim*

Bartz v. Anthropic,
787 F.Supp.3d 1007 (N.D. Cal. 2025) ..... *passim*

Bill Graham v. Dorling Kindersley,
448 F.3d 605 (2d Cir. 2006).....23

Campbell v. Acuff Rose,
510 U.S. 569 (1994)..... *passim*

Clomon v. Jackson,
988 F.2d 1314 (2d Cir. 1993).....25

Estate of Smith v. Cash Money Records, Inc.,
253 F.Supp.3d 737 (S.D.N.Y. 2017).....15

Flo & Eddie v. Sirius XM,
821 F.3d 265 (2d Cir. 2014).....7

Flo & Eddie v. Sirius XM,
No:1-13-cv-5784, Dkt. 51 7

Fox News Network, LLC v. TVEyes, Inc.,
883 F.3d 169 (2d Cir. 2018).....23

Google v. Oracle,
593 U.S. 1 (2021).....17, 21, 23

Hachette Book Grp., Inc. v. Internet Archive,
115 F.4th 163 (2d Cir. 2024) 6, 10, 11, 25

Justice v. Uncharted Labs d/b/a Udio,
No. 25-cv-05026, Dkt. 33 (Sept. 11, 2025 S.D.N.Y.).....22

Kadrey v. Meta Platforms, Inc.,
788 F. Supp. 3d 1026 (N.D. Cal. 2025) ..... *passim*

Kane v. Comedy Partners,
2003 WL 22383387 (S.D.N.Y. Oct. 16, 2003) 15

Monster Commc'ns, Inc. v. Turner Broad. Sys., Inc.,
935 F.Supp. 490 (S.D.N.Y. 1996) 22

NXIVM v. Ross Inst.,
364 F.3d 471 (2d Cir. 2004).....7, 13, 14, 22

Penguin Random House LLC v. Colting,
270 F.Supp.3d 736 (S.D.N.Y. 2017).....16

Perfect 10 v. Amazon,
508 F.3d 1146 (9th Cir. 2007) 6

Romanova v. Amilus,
138 F.4th 104 (2d Cir. 2025) 16

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

Sony Computer Entm't, Inc. v. Connectix Corp.,
203 F.3d 596 (9th Cir. 2000) 6

Thomson Reuters Enter. Ctr. GmbH v. ROSS Intel. Inc.,
193 F.4th 271 (3d Cir. 2026) 13, 16, 21, 24

Twin Peaks Prods., Inc. v. Publications Int'l, Ltd.,
996 F.2d 1366 (2d Cir. 1993).....14

Wright v. Warner Books, Inc.,
953 F.2d 731 (2d Cir. 1991).....21

TABLE OF CITATION CONVENTIONS

Citation	Document
Pls-MSJ.	Class Plaintiffs' Memorandum of Law in Support of Motion for Partial Summary Judgment, Dkt. 1848.
Pls-Opp.	Class Plaintiffs' Memorandum of Law in Opposition to OpenAI Defendants' Motion for Summary Judgment, Dkt. 2262.
Pls-MSFT-Opp.	Class Plaintiffs' Memorandum of Law in Opposition to Microsoft's Motion for Summary Judgment, Dkt. 2260.
OAI-Opp.	OpenAI's Opposition to Class Plaintiffs' Motion for Summary Judgment, Dkt. 2205.
MSFT-Opp.	Defendant Microsoft Corporation's Opposition to Books Plaintiffs' Motion for Partial Summary Judgment, Dkt. 2179.
News-Opp.	News Plaintiffs' Omnibus Memorandum of Law in Opposition to Defendants' Motions for Summary Judgment, Dkt. 2119.
SUF	Class Plaintiffs' Corrected Rule 56.1 Statement of Undisputed Material Facts in Support of Class Plaintiffs' Motion for Partial Summary Judgment, Dkt. 1988.
RSUF	Class and News Plaintiffs' Response to OpenAI's Statement of Undisputed Material Facts in Support of its Motions for Summary Judgment, Dkt. 2202
ASUF	Class Plaintiffs' Counterstatement of Material Facts in Opposition to OpenAI's Motion for Summary Judgment, Dkt. 2247.

INTRODUCTION

Defendants' oppositions are notable in what they do not dispute. As to its downloading, OpenAI does not contest it pirated millions of books—including Plaintiffs' works—from Library Genesis. And OpenAI admits it never trained on many of those same books. As to training, OpenAI does not address its executives' admissions that it was their "research mission" to design commercial products that produce "substitute[s]" for Plaintiffs' works. Pls-MSJ. at 1. Likewise, Defendants do not contest that OpenAI could have used public domain books instead of Plaintiffs' works to train models with "general-purpose" capabilities. *See* Pls-MSJ. at 30-31. As to distributions, Defendants concede that they traded Plaintiffs' works between themselves and to third-parties without permission or payment.

What separates the parties for *Plaintiffs'* summary judgment motion are two legal questions:

First, whether Defendants can retroactively rehabilitate their unauthorized piracy and distributions with a purported training purpose. And *second*, whether the Court must ignore both the substitutive purpose and effect in using Plaintiffs' expression to train book-generating machines. Resolution of either of these legal questions in Plaintiffs' favor compels summary judgment for Plaintiffs. The converse, however, is not true. Even under Defendants' view of the law, bona fide factual disputes preclude summary judgment in their favor. *See* Pls-Opp. at § I-III.

As to the first issue, *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023) ("*Warhol*"), *American Geophysical Union v. Texaco*, 60 F.3d 913 (2d Cir. 1994) ("*Texaco*"), and a long line of similar cases make clear that courts must separately consider uses implicating distinct factual and legal issues under the fair use factors. Defendants' putatively noble downstream purpose cannot alter their objectively substitutive purposes behind pirated initial acquisition or unpaid distributions. Nor can they blind the Court to the distinct harms that

OpenAI's piracy and Defendants' distributions pose—separate from any later use in training. Plaintiffs are thus entitled to summary judgment as to both OpenAI's piracy and Defendants' distributions.

As to the second issue, copyright exists to provide "an economic incentive to create original works" and "the problem of substitution [is] copyright's bête noire." *Warhol,* 598 U.S. at 528, 535. A defendant cannot exploit expressive elements of a copyrighted work without payment and for profit unless they show both that their use is "transformative" (i.e. non-substitutive) *and* that they needed copying to enable the use.

Both are lacking here: Defendants admit that Plaintiffs' works were entirely irrelevant to the "general-purpose" capabilities they rely on to show transformativeness. *See* Pls-MSJ. at 30-31. Defendants instead exploited Plaintiffs' expressive works to improve their models' abilities to substitute for those same works. Defendants' own executives understood the substitution inherent in training their models on Plaintiffs' works. The evidence of GPT-generated substitution [redacted] and in prominent public examples confirms their understanding. Moreover, Defendants undertook this substitutive and unnecessary exploitation of Plaintiffs' works in the face of a developing licensing market to use books in AI-training. Defendants' training use, like their piracy and distribution uses, is not fair.

FACTUAL BACKGROUND

Throughout their oppositions, Defendants incorrectly describe facts as unrebutted, uncontested, or undisputed. While Plaintiffs incorporate their prior filings, Exhibit A identifies the claims in question and the record and briefing citations rebutting them.

ARGUMENT

Plaintiffs seek summary judgment on three distinct uses: (1) OpenAI's initial, objectively substitutive piracy of millions of books, which it stored, catalogued, and made available internally to all employees; (2) OpenAI's use of Plaintiffs' works in training to save money and improve its models' ability to generate substitutive outputs; and (3) Defendants' distribution of Plaintiffs' works as currency in their commercial relationship and to third-party contractors. *Warhol* requires courts to judge the fairness of each use separately. 598 U.S. at 533. Each use's distinct, substitutive purpose under factor one and unique harms under factor four renders them independently not fair.[1]

I. Defendants' Claimed "Ultimate" Purpose Does Not Control Because The Court Must Analyze All Three of Defendants Uses Separately.

Defendants' defense of piracy and distribution amounts to one phrase: it was all training. *See* OAI-Opp. at 8-15; MSFT-Opp. at 9-16. According to Defendants, even when OpenAI *did not use millions of the books it pirated for training*, that too was training. *See* OAI-Opp. at 13-14.

But the law requires courts to analyze distinct uses separately. *See Warhol,* 598 U.S. at 528, 532-33 & nn.4, 7. Fair use "always depends on consideration of the precise facts at hand," *Texaco*, 60 F.3d at 916. Each of Defendants' uses implicates distinct facts and legal principles relevant to the fair use factors. Claimed "ultimate" or downstream purposes, whether of the initial infringer or a third-party, do not control. *See* Pls-MSJ. at 19-22; Pls-Opp. at 4-7, 10-12, 40. Under

[1] Factors two and three—the nature of Plaintiffs' Works and the amount copied—do not vary between uses here; Plaintiffs will rely on their earlier briefing as to both of these factors for each challenged use. Pls-MSJ. at 22-23, 33, 39; Pls-Opp. at 12-13, 30-31; Pls-MSFT-Opp. at 30, 37.

Warhol, even separate quotations in a "critical book review" must stand on their own and a court cannot find fair use simply because "the review, as a whole, serve[s] a different purpose than the book." 598 U.S. at 528 n.4. Indeed, the Government and amici supporting Defendants on training recognize courts must separately analyze uses. *See, e.g.,* DOJ Statement, Dkt. 1682 at 9-10 ("use-by-use approach"); Public Knowledge, Dkt. 2313-1 at 4 ("training separately from data acquisition"); Wikimedia, Dkt. 2361-1 at 6-7 (fair use judged "separately [as] to each such use").

Defendants' claim to a universal "training" purpose is irrelevant and false. *See infra* § II.A; IV.A. OpenAI's initial piracy and Defendants' distributions had unique, substitutive, purposes and effects objectively distinct from each other and from any training use. *See infra* § II; IV; Pls-MSJ. at 5-7, 19-25, 38-39; Pls-Opp. at 7-15, 40.

Defendants offer three responses: first, that the law does not require a "copy-by-copy" analysis; second, that the plaintiffs in this MDL are inconsistent; and third, that training vindicates all their copying, even where the copies were never used in training. All three fail.

A. Distinct Uses Require Distinct Treatment and Defendants' "Copy-By-Copy" Argument Rebuts a Straw-Man.

Defendants argue that a "copy-by-copy" fair use analysis is not required. *See* OAI-Opp. at 10. But Plaintiffs never argued a "copy-by-copy" approach. Plaintiffs have been clear that the proper subject is the separate "use," not each copy. *See, e.g.,* Pls-MSJ. at 19; Pls-Opp. at 4, 6. Copies made during, e.g., tokenization and pre-training may be analyzed as "training" *because* the fair use factors do not meaningfully differ between them. But Defendants' unpaid downloads and distributions implicate distinct issues. Those differences compel separate treatment. *See Warhol*, 598 U.S. at 528, 532-33 & n.4.

Defendants also try to manufacture a difference between Class and News Plaintiffs on separate treatment of separate uses. *See* OAI-Opp. at 14-15; MSFT-Opp. at 12. Wrong. Just as

Class Plaintiffs separately analyzed piracy, training, and distribution, News Plaintiffs separately analyzed training, grounding/RAG, and distribution. *See, e.g.,* News-Opp. at 4 & n.2; *id.* at 35.

B. Defendants' Training Use Cannot Subsume Their Separate Piracy and Distribution Uses.

Defendants claim that "all of the data OpenAI downloaded was for the transformative purpose of training its LLMs," and thus that all infringements constituted a singular "use." *See* OAI-Opp. at 13. While Plaintiffs refute Defendants' claims regarding their "training" purpose below, *see infra* § II.A; IV.A, and while training itself is not fair use, *see infra* § III, *Texaco* makes clear that a later putatively transformative purpose cannot shield distinct infringements from separate fair use scrutiny. *See* 60 F.3d at 924; Pls-MSJ. at 20-22; Pls-Opp. at 9-11.

i. *Texaco* Bars Reliance on Defendants' Claimed Downstream "Training" Purpose to Vindicate Unpaid Downloading or Distributions.

Defendants argue that *Texaco* supports evaluating unpaid acquisition and distributions in light of Defendants' purported "ultimate [transformative] purpose" of "extracting uncopyrightable information" via training. *See* MSFT-Opp. at 15; OAI-Opp. at 12. Not so.

Texaco rejects overriding the substitutive purposes of initial unpaid copying by invoking downstream desires to exploit uncopyrighted facts. *See* 60 F.3d at 924-25. Instead, Texaco's attempt to avoid paying for copies of journal articles—"factor[s] of production" in "efforts to develop profitable products"—was not transformative simply because its scientists wanted the articles for their uncopyrighted facts. *Id.* at 922. "[T]he concept of a 'transformative' use would be extended beyond recognition if it was applied to [] copying simply because [the employee] acted in the course of doing research." *Id.* at 925.

What is true of petroleum engineers is true of software engineers. "You can't just bless yourself by saying I have a research purpose and, therefore, go and take any textbook you want.

That would destroy the academic publishing market if that were the case." *Bartz v. Anthropic*, 787 F.Supp.3d 1007, 1025 (N.D. Cal. 2025) (quoting defense counsel).

ii. Defendants' Remaining Citations Confirm Subsequent Use Cannot Retroactively Justify Pirated Downloading or Distributions.

Campbell v. Acuff Rose, 510 U.S. 569, 592-93 (1994)—which OpenAI cites in support of its "everything is training" approach, OAI-Opp. at 9—refused summary judgment on fair use *because* the challenged song had non-parodic aspects in addition to its transformative parodic aspects. *See* Pls-Opp. at 35. *Campbell* thus demonstrates the required analysis of distinct uses *even within a single song*.

Authors Guild v. Google, Inc., 804 F.3d 202, 217–18 (2d Cir. 2015) ("*Google Books*") did not address initial acquisition because Google lawfully borrowed books which libraries had paid tens of millions of dollars to acquire. *See* Pls-Opp. at 10. Regardless, the court treated both Google's distribution of scans to libraries and its storage of copies *separately* from its search engine. *See* 804 F.3d at 214-30.

Contrary to OpenAI's claims, *see* OAI-Opp. at 12, *A&M Recs., Inc. v. Napster, Inc.*, 239 F.3d 1004, 1014, 1018-19 (9th Cir. 2001) ("*Napster*") rejected fair use for users, including users who downloaded "full, free, and permanent cop[ies]]" to "sampl[e]" them and *later* purchased lawful copies. And *Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 184 (2d Cir. 2024) ("*Internet Archive*") held that the argument that a noble downstream use renders a free download or distribution transformative "would extend the concept of transformativeness 'beyond recognition.'"

Defendants' remaining citations are unpersuasive. *See* OAI-Opp. at 9-15; MSFT-Opp. at 10-16. In *Perfect 10 v. Amazon*, 508 F.3d 1146 (9th Cir. 2007), *Sony Computer Entm't, Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000), and *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d

1510 (9th Cir. 1992) either material was lawfully acquired (leaving no unlawful initial acquisition to address), or the defendant did not possess a copy. *See* Pls-Opp. at 11, 38; *see also Bartz*, 787 F.Supp.3d at 1027-29 (distinguishing).

NXIVM v. Ross Inst., 364 F.3d 471, 475 (2d Cir. 2004) concerned a cult's infringement claim against critical reporting, not initial copying. *NXIVM* only held that bad faith in obtaining an otherwise-unobtainable manuscript did not overwhelm the reports' fair use. *Id.* at 478-479. While the record establishes bad faith, *see* Pls-MSJ. at 5-7, Pls-MSFT-Opp. at 37-38, the requirement for separate analysis is orthogonal to "bad faith." OpenAI's decision to pirate works in preference to its then-on-going lawful acquisition efforts demonstrates both a substitutive purpose and—given the piracy-specific harms from OpenAI's conduct, *see infra* § II.B—a harmful effect distinct from any later training.

Flo & Eddie v. Sirius XM, 821 F.3d 265, 269-70 & n.4 (2d Cir. 2014), purely involved New York state law and, in particular, addressed whether state law provided for an exclusive public performance right for pre-1972 sound recordings. At most the decision holds that, under New York law, a company's internal reproductions of *lawfully purchased* recordings to broadcast them was fair use, where state law also categorically allowed for the uncompensated public performance of sound recordings. *Id.* That has virtually no relevance here, especially because, unlike Defendants, Sirius *bought* or otherwise lawfully acquired the copies it used for broadcasting. *See Flo & Eddie v. Sirius XM,* No:1-13-cv-5784, Dkt. 51, Smith Decl. at ¶ 21-22 (May 30, 2014 S.D.N.Y.).

Kadrey v. Meta Platforms, Inc., 788 F. Supp. 3d 1026 (N.D. Cal. 2025), the lone outlier, misapplied the law on this point, has been questioned by the judge who decided it, *see Kadrey*, No. 23-cv-3417 (N.D. Cal Jul. 8, 2026), Dkt. 767, arose in a decision finding most training is *not*

fair use, and cannot override binding precedent that courts must decide uses on their distinct merits rather than on some downstream purpose. *See supra*; Pls-Opp. at 10-11.

II. Use #1: OpenAI's Free Downloads of Millions of Books From Notorious Pirate Sites Was Infringement as a Matter of Law.

OpenAI infringed when it downloaded for free millions of books it otherwise would have had to purchase or license, stored them on internal servers, [redacted] and made them available to all employees and contractors for any purpose.

Defendants try to distract from the unremarkable proposition that it is not fair use to pirate works a defendant would "ordinarily have to buy." *Napster*, 239 F.3d at 1015. But OpenAI's unpaid acquisition of Plaintiffs' works from pirate libraries to avoid the time and cost of lawful acquisition—depriving Plaintiffs of sales, spurring similar piracy, and saving shadow libraries from "dying"—is not fair use. *See* Pls-MSJ. at 19-25; Pls-Opp. at 7-15. As *Bartz* held, the Court must "contemplate the likely result were the conduct to be condoned as a fair use—namely to steal a work you could otherwise buy (a book, millions of books) so long as you at least loosely intend to make further copies for a purportedly transformative use (writing a book review with excerpts, training LLMs, etc.), without any accountability. As Anthropic itself suggested, 'That would destroy the [entire] publishing market if that were the case.'" 787 F.Supp.3d at 1033. So too here.

A. Factor One: OpenAI's Purpose in Pirating Millions of Books Was Substitutive, Not Transformative.

Later use, even for some purportedly transformative purpose, does not rehabilitate the objectively substitutive purpose of unpaid acquisition. *See supra* § I. That forecloses Defendants' attempt to impute an "ultimate" training purpose to unpaid downloading from shadow libraries. *See* OAI-Opp. at 8-15; MSFT-Opp. at 13-16.

The record refutes Defendants' "ultimate purpose" argument. The undisputed facts show OpenAI pirated for the purpose of obtaining books for free as a speedy alternative to purchasing

or licensing access. *See* Pls-MSJ. at 19-22; Pls-Opp. at 6-11; SUF¶141-42,161-65,169,174-75,182-91,283-304,519; RSUF¶104; *Bartz*, 787 F.Supp.3d at 1015 (Anthropic "preferred to steal [books] to avoid legal/practice/business slog" (cleaned up)). OpenAI even abandoned [redacted] right after it pirated. *See* Pls-Opp. at 8. And OpenAI engaged in precisely the same storage, [redacted], and internal availability which *Bartz* found "objective[ly]" demonstrated a purpose to "substitute for paid copies." 787 F.Supp.3d at 1026; Pls-Opp. at 12.

Underscoring that its downloads were not "for training," OpenAI acquired millions of works it never used in training, yet categorized, organized, and kept them available for any use. *See* Pls-MSJ. at 5-6, 11-12, 22; SUF¶266-67,284.

OpenAI cites *Kadrey* to support its claim that it possessed a training purpose even for works it never trained on and asserts that "[u]nlike in *Bartz*, there is no basis to conclude that OpenAI used the Asserted Works for any purpose other than model training." OAI-Opp. at 14. False. The same acquisition-storage-searching prerequisites which *Bartz*, 787 F.Supp.3d at 1026-28, looked to are present here. And OpenAI made further copies outside training, including via distributions to Microsoft and third-party contractors. *See* Pls-MSJ. at 9-10. Like Anthropic, OpenAI "seems to believe that because *some of* the works it copied were *sometimes* used in training LLMs, [it] was entitled to take for free all the works in the world and keep them forever with no further accounting. There is no carveout, however, from the Copyright Act for AI companies." *Bartz,* 787 F.Supp.3d at 1022 (emphasis added).

OpenAI also claims the [redacted] policies Plaintiffs cite took effect in [redacted]. *See* OAI-Opp. at 13. But OpenAI has pointed to no [redacted]

[redacted] *See* SUF¶247,252-56.

Because precedent bars imputing downstream purposes onto OpenAI's piracy, and because the record shows OpenAI's objectively substitutive purpose in downloading millions of books without payment, factor one favors Plaintiffs.

B. Factor Four: OpenAI's Free Downloads Harmed Plaintiffs' Directly and Would Destroy Their Traditional Markets If "Unrestricted and Widespread."

OpenAI's piracy *by itself* harmed Plaintiffs' (1) through lost sales, (2) by aiding other shadow library users in downloading Plaintiffs' works, (3) by sparking waves of copycat piracy, and (4) by aiding shadow libraries that, before OpenAI's piracy, were "dying." *See* Pls-MSJ. at 23-25; Pls-Opp. at 7, 13-15.

OpenAI—a company valued at approximately one trillion dollars—responds that these harms are "too insubstantial to be counted." OAI-Opp. at 36-37. OpenAI's blithe disregard of the harms its piracy caused is unsurprising, but it is not the law.

First, OpenAI's failure to pay for or license Plaintiffs' works caused lost sales and "plainly displaced demand for [Plaintiffs'] books — copy for copy." *Bartz*, 787 F.Supp.3d at 1033. These harms "deprive[] authors of the revenues due to them as compensation for their unique creations" and alone suffice to find against OpenAI on factor four. *Internet Archive*, 115 F.4th at 195.

Second, the harm from OpenAI's piracy did not stop with its downloads. OpenAI's piracy bolstered the shadow library ecosystem by massively increasing demand for their illicit services. For example, OpenAI's use of Library Genesis led to Anthropic and many others doing the same thing, as well as the creation and distribution of the pirated Books3 set. *See* Pls-Opp. at 13-14. Unrebutted expert testimony shows OpenAI's downloading aided other shadow library users. *Id.* at 14. And shadow libraries attested to being saved from "dying" by OpenAI's piracy. *Id.* at 13-

14. OpenAI's response—that it did not download from the shadow library that posted the "dying" message, *see* OAI-Opp. at 24—proves Plaintiffs' point; even pirate sites OpenAI did not patronize benefitted from the piracy OpenAI set off. The upshot of OpenAI's conduct is that pirate libraries—many of which were otherwise struggling to exist—have been given new life to provide not just the AI industry but also their typical patrons with pirated books.

Thus even if the Court treated piracy as part of "training," the benefits OpenAI's piracy conferred on shadow libraries and their users would weight factor four in Plaintiffs' favor. *See Kadrey*, 788 F.Supp.3d at 1057-58.

Third, the aid OpenAI's piracy rendered shadow libraries—and the copycat piracy OpenAI inaugurated—demonstrate the "decimat[ion] [of] markets for the Works in Suit across formats" which would occur if OpenAI's conduct became (even more) "unrestricted and widespread." *Internet Archive*, 115 F.4th at 195.

Defendants respond that, because they intended to use the books in training, "unrestricted and widespread" conduct of the same sort would not harm Plaintiffs. *See* OAI-Opp. at 36-37. Defendants offer no evidence for this and the record above refutes it. Regardless, Defendants cannot invoke a purported "ultimate purpose" to justify piracy. If it were fair use "to steal … millions of books[] so long as you at least loosely intend" some later use, the market for books would be destroyed. *Bartz*, 787 F.Supp.3d at 1033. Factor four thus weighs strongly for Plaintiffs.

III. Use #2: Fair Use Does Not Protect OpenAI's Substitutive, Commercial, and Unnecessary Use of Plaintiffs' Work to Train LLMs.

OpenAI made unpaid, unlicensed copies of Plaintiffs' books to train machines to imitate and, ultimately, substitute for Plaintiffs' works. OpenAI did so because writing was a [redacted] SUF¶723, and training on Plaintiffs' works rather than public domain alternatives saved money. *See* Pls-MSJ. at 30-32; Pls-Opp. at 21, 25-26; RSUF¶23; SUF¶283-303. Yet Defendants

admit these copies were unnecessary to the "general-purpose" capabilities they invoke as primary evidence of their transformative purpose. *See* Pls-MSJ. at 30-31. The resulting flood of AI-generated works overwhelming Plaintiffs' markets was not only predictable; it was predicted by the very same executives who trained the models. *See* Pls-MSJ. at 13-15. Because that substitutive, unjustified copying threatens the incentives copyright exists to protect, Defendants' use of Plaintiffs' works in training was not fair use.

A. Factor One: OpenAI Used Plaintiffs' Works to Optimize Its Models Towards Generating Competing Output for Cheap.

Factor one asks (1) whether the "purpose and character" of the use is substitutive or transformative, (2) whether the challenged copying possesses a "justification" connecting it to a transformative purpose, and (3) whether the use was commercial. *See* Pls-MSJ. at 19-20, 32-33; Pls-Opp. at 18-19. Because Defendants' use was objectively substitutive and commercial, and because copying Plaintiffs' works was superfluous to Defendants' claimed transformative purpose, factor one weighs strongly for Plaintiffs.

The substitutive capabilities of OpenAI's models derive directly from exploiting protected expression in Plaintiffs' books. *See* Pls-MSJ. at 26-29; Pls-Opp. at 15-16, 19-20, 28-29. Defendants' oppositions ignore that OpenAI's executives, including Jack Clark and Tarun Gogineni, repeatedly confirmed OpenAI's models would "substitute" for human authors and targeted Plaintiffs in this case for substitution as their "research mission." Pls-MSJ. at 13-15. The purpose of Defendants' use *of Plaintiffs' Works in particular* was thus avowedly substitutive, not transformative.

As to justification, OpenAI's testing confirmed it could train "general purpose" models on public domain books—at the cost of higher training spend and reduced ability to generate competing book-like text. *See* Pls-MSJ. at 30-32; Pls-Opp. at 21. Defendants' experts agree that

copyrighted books improved models' generation of competing, book-like text while being unnecessary to any of their models' general-purpose capabilities. *See* Pls-MSJ. at 27, 30-31; SUF¶745-46,748. Using Plaintiffs' works was thus unconnected to any transformative end and served to improve Defendants' models substitutive character (and save money) in service of their commercial gain. Factor one weighs decisively in Plaintiffs' favor.

Defendants respond that the law bars considering objective evidence of their substitutive, money-saving purpose, and that their true purpose was training a general-purpose LLM. *See* OAI-Opp. at 16-17; MSFT-Opp. at 18-21. Both arguments fail.

i. Factor One Favors Plaintiffs Because Defendants' Purpose Was Objectively Substitutive.

The "problem of substitution" is the focus of the transformativeness inquiry and, where the use risks substitution, any transformative character "must be balanced against the commercial nature of the use." *Warhol*, 598 U.S. at 528, 532-33. The record unequivocally shows Defendants used Plaintiffs' works in training precisely to improve their models' abilities to substitute for Plaintiffs' works for Defendants' commercial gain (and to save compute costs). *See* Pls-MSJ. at 10-15, 31; Pls-Opp. at § II.A.a, II.A.c.

That record of substitutive purpose—and Defendants actions to achieve it—distinguish this case from *Kadrey* and *Bartz,* whose factor one holdings rested on records lacking such evidence. *See* Pls-Opp. at 27-28. *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intel. Inc.,* 193 F.4th 271, 283-85 (3d Cir. 2026) ("*Ross"*)—following *Warhol*—shows that factor one focuses on the degree to which even a non-infringing end product aims to compete with the original. *See* Pls-Opp. at 2-3, 16-21.

While Defendants downplay *Ross* as not about generative models, *see* OAI-Opp. at 20, the same is true of every appellate decision they cite and *Ross* is incontestably a closer match than,

e.g., *Google Books*. Regardless, *Ross* refutes Defendants' core argument: that exploiting protected expression to build a competing commercial product is "transformative" rather than substitutive.

Defendants' contrary citations, *see* OAI-Opp. at 16-21, represent Defendants misreading of *Campbell*, 510 U.S. at 590-93. *Campbell* held only that parody or similar uses with a "critical aspect" toward the original can be transformative rather than substitutive and it rejected using expression simply to "avoid the drudgery in working up something fresh." *Id.* at 580. And Defendants do not contest that their models render detailed summaries of Plaintiffs' works—[redacted] SUF¶711—which the Second Circuit has found not transformative. *See Twin Peaks Prods., Inc. v. Publications Int'l, Ltd.,* 996 F.2d 1366, 1376 (2d Cir. 1993).

To be clear: Plaintiffs are not arguing that commentary is not fair use, Plaintiffs are arguing that OpenAI's training use *is not commentary* (or any use possessing a "critical aspect" toward Plaintiffs' originals). Defendants used Plaintiffs' works because they made their products better at competing against Plaintiffs and because they saved money. *See supra* § III. That is the purpose of "avoid[ing] the drudgery"—and expense—"in working up something fresh" that *Campbell* rejects. 510 U.S. at 580.

Defendants' new citations, *see* OAI-Opp. at 20, confirm their misapplication of *Campbell*:

- *Arica Inst., Inc. v. Palmer*, 970 F.2d 1067, 1077-78 (2d Cir. 1992)—reflects *Campbell*'s protection of commentary and relied on a strong "presumption of fairness" which *Warhol*, 598 U.S. at 537 n.7, rejected. *Arica's* distinction "between the infringing portions of defendant's work and the work as a whole" reflected that, unlike Defendants' exploitation of Plaintiffs' expression here, the original portions

of the challenged work did not derive from use of the *Arica* plaintiff's expression. *Arica*, 970 F.2d at 1078.

- *Kane v. Comedy Partners*, 2003 WL 22383387 at *4, (S.D.N.Y. Oct. 16, 2003) and *Estate of Smith v. Cash Money Records, Inc.*, 253 F.Supp.3d 737, 750 (S.D.N.Y. 2017) straightforwardly apply *Campbell* to clips used to "critically examine" the plaintiff's work (*Kane*) and a rap song's transformative sampling (*Smith*). Neither applies to Defendants' non-critical use.

Plaintiffs already addressed Defendants' "intermediate copying" citations. *See* Pls-MSJ. at 27-29; Pls-Opp. at 28-30. Briefly, these cases only demonstrate that copying functional elements for a purpose divorced from a work's protected, creative expression (and where copying was necessary to the use) can be transformative. For example, neither *Google Books*, 804 F.3d at 206, which "test[ed] the boundaries of fair use," nor *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014) involved exploiting authors' expression to build a substitutive commercial product. The uses at most gave the public information about the books. By contrast, Defendants unnecessarily exploited the expressive aspects of Plaintiffs' books to improve their products' abilities to flood the market with competing works. Such uses are neither transformative nor fair.

ii. Factor One Favors Plaintiffs Because Defendants Lack Justification for Their Use of Plaintiffs' Works.

Factor one requires a justification connecting Defendants' use of Plaintiffs' works to their purportedly transformative end. *See Warhol*, 598 U.S. at 547 & n.21. But Defendants concede that they did not need Plaintiffs' works—or any copyrighted books—to create general-purpose LLMs, and that training on Plaintiffs' books do not enable their models' general-purpose capabilities in any legally relevant sense. *See* Pls-MSJ. at 29-32; Pls-Opp. at 25-26. [redacted]

[redacted]

[redacted] *See* Pls-MSJ. at 31-32; Pls-Opp. at 17, 21, 26. That undisputed lack of justification sets factor one in Plaintiffs' favor.

Defendants respond that, under *Warhol*, they need no meaningful justification because their copying was not substitutive. *See* OAI-Opp. at 17-20. That simply rehashes the substitution argument above. *See supra* § III.A.i. *Warhol* itself held that uses implicating substitution—like Defendants'—require special justification. *See* 598 U.S. at 532. And no precedent constrains substitution to direct reproduction, particularly where, as here, the substitutive character derives from exploiting Plaintiffs' expression for commercial gain. *See Ross*, 193 F.4th at 283-85.

Defendants also argue their copying was justified because their LLMs perform many non-substitutive functions. *See* OAI-Opp. at 16. But justification tests whether even a "transformative" use actually needed to use copyrighted material and bars copying out of "ease" rather than "necessity." *Ross*, 193 F.4th at 286. A bootleg DVD is not a justified transformative use simply because the vendor also included an original review or a list of the digits of pi. Nor may Defendants optimize a book-generating machine simply because it also debugs software or recommends restaurants in lower Manhattan. *See Penguin Random House LLC v. Colting*, 270 F.Supp.3d 736, 751 (S.D.N.Y. 2017) ("tacking on these few pages [of criticism] does not provide safe harbor for an otherwise infringing work").

Nor do categories of justification discussed in *dicta* in *Romanova v. Amilus*, 138 F.4th 104, 114-19 (2d Cir. 2025) apply to Defendants' training use. In each case *Romanova* cited, copying was needed to achieve the end offered as justification. Google's "ngrams" feature, for example, could not chart historical word frequencies without scanned books. *See id.* at 116. By contrast, as Defendants admit, every part of Defendants' service could function had they not copied a single copyrighted book—much less Plaintiffs' works. *See* Pls-MSJ. at 29-32; Pls-Opp. at 20-21, 25-26.

Google v. Oracle, 593 U.S. 1, 30 (2021) found Google's use justified *because* it copied functional material "*only insofar as needed* to allow programmers to call upon [tasks implemented by the copyrighted code]." As OpenAI's research showed (and Defendants' experts concede), OpenAI did not copy "only insofar as needed," *id.*, to enable general-purpose capabilities. It copied Plaintiffs' works to save money and improve performance on a specific, substitutive, "top use case[]." SUF¶723; Pls-Opp. at 20. Defendants' inability to show justification sets factor one in Plaintiffs' favor.

iii. At Minimum, Factual Disputes Regarding the Nature of LLM Training Precludes Summary Judgment in Defendants' Favor.

Plaintiffs thus prevail on factor one and, because Defendants' remaining arguments depend on training being a justified transformative use, the Court should grant Plaintiffs summary judgment on training.

Even on Defendants' view of factor one, the Court would still deny Defendants' motions because their factual claims about the training process are genuinely disputed. As Plaintiffs' Opposition explains, Defendants' employees and executives admitted the training objective is not understanding "unprotected facts" but [redacted] *See* Pls-Opp. at 17, 22-23.

Defendants again emphasize anti-regurgitation measures as showing their goal in training was not to replicate Plaintiffs' expression. *See* OAI-Opp. at 18-19. Contemporaneous statements [redacted] contradict Defendants' claims. *See* Pls-Opp. at 23-25; RSUF¶73,75. More importantly, Defendants do not contest [redacted]. *See* Pls-MSJ. at 13; Pls-

Opp. at 24; SUF¶564; RSUF¶65. Thus, even on Defendants' view, Plaintiffs prevail as to those models.

B. Factor Four: Displacement Caused By The Flood of AI-Generated Works and Harms to the Already-Existing AI Training Licensing Market Weigh Strongly Against Fair Use For Training.

Defendants' uncompensated use of Plaintiffs' works to improve models' abilities at generating substitutive outputs harmed Plaintiffs' traditional markets though a flood of competing works and deprived them of fees from the AI-training licensing market. *See* Pls-MSJ. at § III.C; Pls-Opp. at § II.C. Defendants' contrary arguments fail factually and legally.

i. The Flood of AI-Generated Works Harms Plaintiffs.

Factually, the flood of AI-generated works harms the traditional book market. Plaintiffs' testimony, uncontested [redacted] data, expert analysis, and Defendants' admissions confirm that models trained on Plaintiffs' works generate books in minutes for less than a dollar, and that those books are winning sales, recognition, and prizes in the traditional book market. *See* Pls-Opp. at 33-34; SUF¶845-57. Like Plaintiffs, the working authors who filed as amici illustrate the "emotional [and] financial" harms of Defendants' training use and their impacts on the incentive to create. *See* Authors *et al*, Dkt. 2402-2 at 10-14; *id.* at 10 ("[m]y sales have dropped dramatically in the last two years"); *id.* at 14 ("productivity and even my ability to write have been very adversely affected"). Defendants at most dispute the magnitude of that impact; they do not carry their burden to show a lack of market harm. *See* Pls-MSJ. at 33; Pls-Opp. at 31.

Defendants principally rely on expert analysis which ignores [redacted] *See* OAI-Opp. at 6, 30-31; Pls-Opp. at 33. [redacted] data conclusively refutes Defendants' no-impact claims: [redacted]. *See*

SUF¶845,851-54; RSUF¶125-131. As Plaintiffs' expert Michael Lasinski showed, correcting Defendants' omissions reveals Plaintiffs' sales declines. *See* Pls-Opp. at 33.

Defendants' incorrectly claim that Plaintiffs are too famous for GPT-generated works to harm them. *See* OAI-Opp. at 31-32. Plaintiffs' testimony and prominent examples show notoriety, far from insulating, makes Plaintiffs targets for AI-generated works seeking to impersonate and cannibalize sales. *See* Pls-Opp. at 34; ASUF¶27; SUF¶873-97. Dr. Shan showed Defendants' models generate competing sequels and summaries, and [redacted] *See* Pls-Opp. at 25. AI-generated works winning literary prizes and publishing deals—as well as Mr. Lasinski's analysis showing sales declines post-ChatGPT—demonstrate that the markers of quality and distribution channels, which Defendants posit as barriers to GPT-generated substitution, are illusory. *See* Pls-Opp. at 33.

Nor do Defendants address the impact of "unrestricted and widespread" training on Plaintiffs' works. Dr. Shan's unrebutted analysis shows models' capabilities to produce substitutive output has risen with every release and will only keep rising. *See* Pls-Opp. at 33-34. Thus, even beyond the impact of the challenged models, unrestricted and widespread conduct like Defendants' would inevitably displace Plaintiffs' works.

ii. The Market to License Books for AI-Training Exists And Is Rapidly Expanding.

The evidence of an active market to license books for AI training is overwhelming; Defendants themselves have signed such licenses.

Defendants' respond that none of the more than 120 licenses worth more than $1 billion signed to date, SUF¶775-840, is comparable to their conduct. *See* OAI-Opp. at 35-36; MSFT-Opp. at 33-34. But Mr. Lasinski's comparability analysis identifies licenses covering precisely the conduct challenged here: using books to train AI models capable of generating substitutive outputs.

See Lasinski Decl. Ex. B ¶49–138. And, based on past successful licensing regimes, the market to license books for training possesses all pre-requisites—e.g. intermediaries, motivated participants, identifiable rightsholders—for continued expansion. ASUF¶52-56. Amici confirm that the licensing market for AI-training, including for books, is growing rapidly. *See, e.g.,* American Assn. of Publishers *et al*, Dkt. 2304-1 at 14-17; Copyright Alliance, Dkt. 2359-1 at 5-7; AI Coalition, Dkt. 2386-1 at 12-16; Open Markets Institute, Dkt. 2372-1 at 13-19; RIAA *et al*, Dkt. 2393-1 at 13-16; Human Consent Foundation *et al*, Dkt. 2417-1 at 12-18; AI Startups, Dkt. 2460-1 at 11-20; SAG-AFTRA et al, Dkt. 2332-1 at 22.

A prime example of a comparable license is Microsoft's agreement with [redacted] for AI training. RSUF¶121, SUF¶790–792,814,825; Lasinski Decl. Ex. B ¶49–138. Defendants' main response is that they consider the agreement a failure. *See* MSFT-Opp. at 34. But the existence of buyer's remorse necessarily demonstrates the existence of a market to buy from. Whatever Microsoft says about the deal now—i.e. after being sued for failing to license similar material—cannot change the fact that Microsoft was able to enter into such a license in the first place, thus demonstrating that the market for such licenses both exists and is likely to develop. *See Texaco,* 60 F.3d at 930. In any event, Defendants' contemporaneous and post-licensing statements show their made-for-litigation regret is false. *See* Pls-MSJ. at 12, 37-38; Pls-Opp. at 32

iii. The Court Must Consider Harms from Market Displacement As a Matter of Law.

Defendants argue that factor four only considers harms from reproduction and display of copyrighted expression to the public. *See* OAI-Opp. at 26-30; MSFT-Opp. at 28-31. But that confuses the legal test for infringement with the factual question of what "effect" Defendants' use had on the market for Plaintiffs' works. *See* Pls-Opp. at 36. As shown above, the effect is seismic.

Defendants again rely on the *Campbell*-line of "criticism as fair use" cases. But those cases hold only that owners cannot weaponize copyright against critics or wags. None approve of ignoring "an obvious way that a new technology might severely harm the incentive to create." *Kadrey*, 788 F.Supp.3d at 1055. *Campbell* itself recognized harm from a use's non-parodic character. *See* 510 U.S. at 591-92. *Texaco*, 60 F.3d at 924-25, similarly shows harms are cognizable even where copies are made for internal research and not shown to the public. And *Ross*, 193 F.4th at 288-89, found the threat of substitution from a competing legal research tool trained on the plaintiffs' materials weighted factor four in the plaintiff's favor, even though the tool displayed zero copyrighted content to users.

As to *Google Books* and *Oracle*, neither involved a use which generated competing works or competed with the originals in their market. *Google Books*, 804 F.3d at 222-24, found the demand satisfied by unprotected material was insufficient for displacement, and *Oracle*, 593 U.S. at 36-38, involved functional material in a different market. The opposite is true for Defendants' exploitation of Plaintiffs' expression here.

Defendants' other citations do not contravene well-established law that the Court must consider harms to Plaintiffs' markets from substitution posed by Defendants' non-critical use:

- *Wright v. Warner Books, Inc*., 953 F.2d 731, 739 (2d Cir. 1991) addressed "factual content" in a biography (*Campbell*'s heartland), applied a presumption of fairness which *Warhol,* 598 U.S. at 537 n.7, rejected, and found effects from exploiting plaintiff's expression "unlikely." Here, Defendants' use exploited Plaintiffs' expressive choices, and harm is already occurring.
- *Arica*, 970 F.2d at 1078-79 found that "marginal amounts of expressive content [] taken from [plaintiff's] works" would have a "negligible effect on the market." But

Defendants here did not exploit a "marginal amount" and the effect is not "negligible." As *Arica* recognized, while Defendants are permitted to create "competing work[s]," "[t]he copyright statute [] constrains [them] from using plaintiff's original expression in doing so." *Id.* at 1078.

- *Monster Commc'ns, Inc. v. Turner Broad. Sys., Inc.*, 935 F.Supp. 490, 495–96 (S.D.N.Y. 1996) is another, inapplicable *Campbell*-derived criticism case.
- 4 Patry on Copyright § 10:150, 10.155.40 ("*Patry*"): Defendants cite *Patry* as neutral legal authority, but Mr. Patry represents AI companies in copyright litigation over these same issues. *See Justice v. Uncharted Labs d/b/a Udio,* No. 25-cv-05026, Dkt. 33 (Sept. 11, 2025 S.D.N.Y.).[2] Regardless, Section 10:150 agrees harms "caused by the use of the expression" are cognizable, including from, e.g., detailed summaries or unauthorized sequels. Both are present here. *See* Pls-Opp. at 24-25. As to Section 10:155.40, and among other errors, it labels "market dilution" a "never adopted[] theory," even though *Kadrey* adopted it months earlier.

Defendants argue harms from GPT-generated books are too "far removed" from training to count. OAI-Opp. at 29-30. Not so. Defendants' contemporaneous statements and evaluations affirmed that, by training on books, outputs from GPT models would predictably substitute on "Amazon" for the "personal" and "deeply practiced labor" of authors. SUF¶750.

Defendants' causation argument is also legally wrong. Defendants' primary authority, *NXIVM,* 364 F.3d at 481-82, is a straightforward application of *Campbell*'s protection of criticism. *See supra* § III.A.i. *Google Books*, 804 F.3d at 223-24, expressly *did* consider displacement from

[2] Similarly, Public Knowledge, an amicus supporting Defendants, *see* Dkt. 2313-1, lists each Defendant in its highest "$100,000 and above" donor tier. *See* "Funding Sources," *Public Knowledge,* https://publicknowledge.org/sources-of-funding-for-public-knowledge/.

snippet view's exposure of unprotected material but found it factually too small. And *Oracle*, 593 U.S. at 35-38, turned on the functional nature of code and that Google's use did not compete in Oracle's market. Here Defendants' use has no critical bearing on Plaintiffs' original expression and the impact is seismic. This is not a case where a work "inspires" a human reader to write something of their own. Defendants exploited Plaintiffs' expression as a "factor of production," *Texaco,* 60 F.3d at 922, and to optimize their book-generating machine on the cheap. The books it generates—competing with and drawing sales from Plaintiffs' works—are plainly "effects" under factor four.

iv. The Court Must Consider Harms from Lost Licensing Fees as a Matter of Law.

Defendants' argument that lost licensing fees are not cognizable similarly fails. Because Defendants' use was substitutive—and any transformative character bears no relationship to the challenged copying—lost licensing revenues are a cognizable harm. *See supra* § III.A; *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 179-81 (2d Cir. 2018) (licensing market for transformative use cognizable given threat of substitution). *Campbell* and Defendants' other citations concerned licensing markets unlikely to develop given the critical use (e.g. *Campbell*'s parody, *Bill Graham v. Dorling Kindersley*, 448 F.3d 605, 609-10 (2d Cir. 2006)'s biography/history, and *NXIVM*'s critique) or rested on findings that no comparable licensing activity existed (e.g. *Google Books*). Here, Defendants' use has no critical bearing and a comparable licensing market exists. *See Texaco*, 60 F.3d at 931 (rejecting circularity because "there [was] a ready market").

Nor does it matter that Plaintiffs have not entered such licenses. *See* OAI-Opp. at 34. The question is the existence or "likely development" of the market generally. *See Texaco,* 60 F.3d at

930-31. A plaintiff's non-participation does not refute existence of the market. *See* Pls-Opp. at 32. Regardless, Plaintiffs repeatedly affirmed their willingness to license. *See id*.

Defendants argue courts find "no licensing-market harm" even where licenses exist. *See* OAI-Opp. at 35-36. But the two cites—*HathiTrust*, 755 F.3d at 100 and *Google Books*, 804 F.3d at 226-27—relied on records lacking the licenses for comparable conduct here. *See supra* § III.B.ii.

Bartz and *Kadrey*'s licensing holdings similarly rested on findings regarding the defendants' conduct and the market at the time which do not apply here. *See* Pls-MSJ. at 28; Pls-Opp. at 35-36. Instead, *Ross*, 193 F.4th at 287-88, shows the Court must consider the licensing market where, as here, the use has a competitive and substitutive effect on Plaintiffs' original markets.

IV. Use #3: Defendants' Unpaid "Horse-Trading" Distributions of Plaintiffs' Works Were Not Fair Use.

A. Distribution is a Separate Use From Initial Acquisition or Training And is, Independently, Not Fair.

The law requiring separate treatment of piracy and training applies equally to Defendants' distributions. And, as with piracy, the record contradicts Defendants' claims that distributions were "for training." *See* OAI-Opp. at 7, 13-14; MSFT-Opp. at 22-23.

First, neither of OpenAI's transfers of [redacted] training data was for training [redacted] *See* SUF¶105,573-77. Contrary to OpenAI's misrepresentations, *see* OAI-Opp. at 7 & n.4, the first transfer included [redacted]. *See* Pls-MSJ. at 9; SUF¶573-77. And the second was part of a non-transformative [redacted] with Microsoft. *See* Pls-MSJ. at 9, 38-39.

Second, OpenAI's distribution to third party contractors [redacted] was not training; it was, almost literally, *Texaco*, 60 F.3d at 924-25. OAI-

Opp. at 7-8. The principal difference is that OpenAI, unlike Texaco, did not even pay for its initial copy. *See* Pls-MSJ. at 38-39.

Third, Microsoft's distributions, were part of the same [redacted] not training, discussed above. *See* Pls-MSFT-Opp. at 10-15; 23-33.

As to factor four, distributing copyrighted works to third parties harms the market and is not fair use *even if* it "facilitate[d]" downstream training. *See Internet Archive*, 115 F.4th at 184; Pls-MSJ. at 38-39; Pls-MSFT-Opp. at 30-33, 37-38.

B. OpenAI's Download of [redacted] Distributions to Contractors, and [redacted] Are Infringing Conduct At Issue In This Case.

OpenAI argues that [redacted] its distributions to contractors, and [redacted] go "beyond the In-Scope models." *See* OAI-Opp. at 37-38. OpenAI's argument ignores that the Court's Order regarding models-in-suit also denied OpenAI's motion to strike Plaintiffs' "downloading" claim, held that Plaintiffs pled direct infringement, and stated that Plaintiffs were not "lock[ed]" into particular "legal theor[ies]" *regardless* of whether OpenAI's infringements connected to an "ultimate purpose of . . . train[ing] OpenAI's LLMs." *See* Dkt. 707 at 1, 3–5 & n.2. Regardless, Plaintiffs' opposition explains why the challenged distributions are in this case, *see* Pls-Opp. at 40, and OpenAI's argument represents another unavailing attempt to recast all its uses as training, *but see supra* § I.

OpenAI's argument is also waived. The parties litigated discovery motions which OpenAI either lost or, after Plaintiffs moved, agreed to produce. *See* Dkt. 848 at 5, 8 (denying protective order for 30(b)(6) topics including Books3); Dkt. 950 at 6 (memorializing agreement regarding book summarization materials); Dkt. 720 (9/25/2025 Hearing Tr.) at 128:19-133:17 (agreeing to produce documents regarding "Bing Index data received from Microsoft"). And even on OpenAI's view, the court can conform pleadings to the summary judgment record. *See Clomon v. Jackson,*

988 F.2d 1314, 1323 (2d Cir. 1993). Because the record is fully developed and deferring would be massively inefficient, the Court should decide these issues now.

V. Microsoft is Vicariously Liable for OpenAI's Infringement.

Microsoft's opposition on vicarious liability rests on claims regarding (1) its rights under the parties' agreements and (2) its financial benefits from OpenAI's infringement. MSFT-Opp. at 37-42. But the agreements clearly give Microsoft power over relevant and infringing acts by OpenAI. *See* Pls-MSJ at 8, 40. And Microsoft benefitted financially from using Plaintiffs' works in training, since it reduced costs and improved models' capabilities on a key use case Microsoft itself marketed. *See* Pls-MSJ at 40; SUF¶723,740-44.

CONCLUSION

The Court should grant Plaintiffs' motion for partial summary judgment.

Dated: November 20, 2026

/s/ Justin A. Nelson
Justin A. Nelson
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, TX 77002
Telephone: 713.651.9366
jnelson@susmangodfrey.com
Interim Lead Counsel for Class Plaintiffs

CERTIFICATE OF WORD COUNT

In accordance with Local Civil Rule 7.1(c) and the Court's Order at Dkt. 1664, 25-md-3143, I certify that the foregoing Reply Memorandum of Law is 7,494 words, including footnotes but exclusive of the caption page, table of contents, table of authorities, certificates and signature block. The basis of my knowledge is the word count feature of the word-processing system used to prepare this memorandum.

/s/J. Craig Smyser
Signature

CERTIFICATE OF SERVICE

I hereby certify that on November 20, 2026, I caused the foregoing document to be electronically filed with the Clerk of the United States District Court for the Southern District of New York using the CM/ECF system, which shall send electronic notification to all counsel of record.

/s/J. Craig Smyser
(Signature)