UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE: OPENAI, INC., COPYRIGHT INFRINGEMENT LITIGATION This document relates to: 23-cv-8292 23-cv-10211 24-cv-84 25-cv-3297 25-cv-3482 25-cv-3483	25-md-3143 (SHS) (OTW) Hon. Sidney H. Stein Hon. Ona T. Wang.

MEMORANDUM OF LAW IN SUPPORT OF OPENAI DEFENDANTS' OBJECTION TO DISCOVERY ORDER AT ECF NO. 846 PURSUANT TO FEDERAL RULE OF CIVIL PROCEDURE 72(A)

TABLE OF CONTENTS

Page

I. INTRODUCTION 1
II. BACKGROUND 4
A. Class Plaintiffs Allege Copyright Infringement 4
B. OpenAI Asserts Attorney-Client Privilege Over Its Reasons for Deleting the Datasets. 6
C. Magistrate Judge Wang Orders OpenAI to Disclose *All* Communications Relating to Reasons for Deletion 8
III. STANDARD OF REVIEW 9
IV. ARGUMENT 9
A. OpenAI's Reasons for Deleting the Datasets Are Privileged 9
B. OpenAI Did Not Waive Privilege 11
1. OpenAI Did Not Waive Privilege by Partial Disclosure 11
2. OpenAI Did Not Waive Privilege by Changing Positions 17
3. OpenAI Did Not Waive Privilege by Putting Privileged Communications at Issue 20
V. CONCLUSION 25

TABLE OF AUTHORITIES

Page

CASES

AFP Imaging Corp. v. Philips Medizin Sys.,
1993 WL 541194, at *3 (S.D.N.Y. Dec. 28, 1993) 3

Aiossa v. Bank of Am., 2011 WL 4026902 (E.D.N.Y. Sept. 12, 2011) 14

Arista Records, LLC v. Lime Grp., LLC, 2011 WL 1642434 (S.D.N.Y. Apr. 20, 2011) 24, 25

Briese Lichttechnik Vertriebs GmbH v. Langton, 272 F.R.D. 369 (S.D.N.Y. 2011).....16

Bryant v. Media Right Prods., Inc., 603 F.3d 135 (2d Cir. 2010) 22

Catskill Dev., L.L.C. v. Park Place Entm't Corp., 206 F.R.D. 78 (S.D.N.Y. 2002) 9

Chase Manhattan Bank N.A. v. Drysdale Secs. Corp., 587 F. Supp. 57 (S.D.N.Y.1984).....4, 23

Chelsea Hotel Owner LLC v. City of New York,
2024 WL 323216 (S.D.N.Y. Jan. 29, 2024),
aff'd, 2025 WL 880530 (S.D.N.Y. Mar. 21, 2025).....19

Chin v. Rogoff & Co., 2008 WL 2073934 (S.D.N.Y. May 8, 2008) 20

Digital-Vending Servs. Int'l, LLC v. Univ. of Phoenix Inc.,
2010 WL 11450510 (E.D. Va. Apr. 22, 2010) 19

E.E.O.C. v. George Washington Univ., 502 F. Supp. 3d 62 (D.D.C. 2020).....15

Fitzgerald Publ'g Co. v. Baylor Publ'g Co., 807 F.2d 1110 (2d Cir.1986).....22

Gardner v. Major Auto. Cos., Inc., 2014 WL 1330961 (E.D.N.Y. Mar. 31, 2014) 3, 23

Green v. Beer, 2010 WL 3422723 (S.D.N.Y. Aug. 24, 2010).....22

Halo Elecs., Inc. v. Pulse Elecs., Inc., 579 U.S. 93 (2016).....25

In re County of Erie, 473 F.3d 413 (2d Cir. 2007) (*Erie I*) 10

In re County of Erie, 546 F.3d 222 (2d Cir. 2008) (*Erie II*) ..... *passim*

In re DG Acquisition Corp., 151 F.3d 75 (2d Cir. 1998).....18

In re FirstEnergy Corp., 154 F.4th 431 (6th Cir. 2025) 14, 15

In re Grand Jury Proc., 219 F.3d 175 (2d Cir. 2000).....11, 16

Page

Cases—continued:

In re Honeywell Int'l, Inc. Sec. Litig., 230 F.R.D. 293 (S.D.N.Y. 2003).....3, 17, 18, 19

In re Keurig Green Mountain Single-Serve Coffee Antitrust Litig.,
2020 WL 3034792 (S.D.N.Y. June 5, 2020) 17

John Doe Co. v. United States, 350 F.3d 299 (2d Cir. 2003) 11

Kovacs v. Hershey Co., 2006 WL 3054167 (D. Colo. Oct. 25, 2006).....15

Laborers Local 17 Health Benefit Fund v. Philip Morris, Inc.,
1998 WL 414933 (S.D.N.Y. July 23, 1998) 22

Leviton Mfg. Co. v. Greenberg Traurig LLP, 2010 WL 4983183 (S.D.N.Y. Dec. 6, 2010),
objections overruled, 2011 WL 2946380 (S.D.N.Y. July 14, 2011) 24, 25

MacNamara v. City of New York, 249 F.R.D. 70 (S.D.N.Y. 2008).....10

Murtha v. New York State Gaming Comm'n, 2020 WL 5504311 (S.D.N.Y. Sept. 9, 2020).....14

Maiurano v. Cantor Fitzgerald Sec., 2021 WL 4993060 (S.D.N.Y. Oct. 27, 2021).....14

Nat'l Football League v. PrimeTime 24 Joint Venture,
131 F. Supp. 2d 458 (S.D.N.Y. 2001).....22

Okla. Firefighters Pension & Ret. Sys. v. Musk,
2025 WL 2982003 (S.D.N.Y. Oct. 23, 2025) 23

RMED Int'l, Inc. v. Sloan's Supermarkets, Inc., 2003 WL 41996 (S.D.N.Y. Jan. 6, 2003) 18

Rhone-Poulenc Rorer Inc. v. Home Indem. Co., 32 F.3d 851 (3d Cir. 1994) 22

S.E.C. v. Yorkville Advisors, LLC, 300 F.R.D. 152 (S.D.N.Y. 2014).....19

Sanofi-Synthelabo v. Apotex Inc., 299 F. Supp. 2d 303 (S.D.N.Y. 2004),
on reconsideration, 363 F. Supp. 2d 592 (S.D.N.Y. 2005) 11

Scott v. Chipotle Mexican Grill, Inc., 67 F. Supp. 3d 607 (S.D.N.Y. 2014) 23

Sekisui Am. Corp. v. Hart, 945 F. Supp. 2d 494 (S.D.N.Y. 2013) 9

Standard Chartered Bank PLC v. Ayala Int'l Holdings (U.S.) Inc.,
111 F.R.D. 76 (S.D.N.Y.1986) 23

Sw. Airlines Co. v. Saxon, 596 U.S. 450 (2022) 13

Page

Cases—continued:

Swift Spindrift, Ltd. v. Alvada Ins., Inc., 2013 WL 3815970 (S.D.N.Y. July 24, 2013).....17

Twin Peaks Prods., Inc. v. Publ'ns Int'l, Ltd., 996 F.2d 1366, 1382 (2d Cir. 1993).....22

United States v. Bilzerian, 926 F.2d 1285 (2d Cir.1991).....16, 23

United States v. Mejia, 655 F.3d 126 (2d Cir. 2011).....10

Upjohn Co. v. United States, 449 U.S. 383 (1981).....10

STATUTE AND RULES

17 U.S.C. § 504.....21

Federal Rule of Civil Procedure 72 9

Federal Rule of Evidence 502.....2, 11, 15, 16, 17

OTHER AUTHORITIES

Due To, *The Am. Heritage Coll. Dictionary* (5th ed. 2022) 13

24 Wright & Miller, Fed. Prac. & Proc. Evid. § 5507.2 (1st ed. 2025 update) 18

I. INTRODUCTION

OpenAI respectfully objects to Magistrate Judge Wang's November 24, 2025 Opinion and Order (the "Order") (ECF 846). The Order compels production of OpenAI's attorney-client privileged communications and directs depositions of OpenAI's in-house counsel on privileged matters. The Order justifies that extraordinary intrusion into OpenAI's attorney-client relationships on the theory that OpenAI impliedly waived privilege. Order at 24. But the Order reached that conclusion only by dispensing with controlling precedent, dramatically expanding waiver doctrine, and misconstruing the record—all in the service of permitting the plaintiffs to access attorney-client communications that are irrelevant to the copyright-infringement claims in this case.

The class plaintiffs allege that OpenAI infringed copyrights in their books by downloading unauthorized copies and using them to train OpenAI's artificial-intelligence systems. The plaintiffs have been granted extensive discovery into OpenAI's training data, when and how it was acquired, and how OpenAI used it. But instead of building their case with that evidence, plaintiffs turned their attention to an extraneous issue: OpenAI's 2022 decision to delete two datasets—Books1 and Books2—created by former OpenAI researchers and briefly used to train two discontinued GPT models. OpenAI's *deletion* of the datasets has nothing to do with OpenAI's alleged infringing *use* of the plaintiffs' works to train AI systems years earlier. Plaintiffs have nevertheless benefited from access to recovered copies of Books1 and Books2 along with discovery into when and how the data used to produce the datasets was downloaded, who downloaded it, the timing and factual subject matter of discussions about deleting the datasets, and the technical steps behind deletion.[1] Dissatisfied with the results of that broad discovery, the class

[1] *See* Ex. A (July 25, 2025 Trinh Dep. Tr.) 19:12–20:18, 70:25–72:3, 170:19–171:9 (discussing when the datasets were downloaded, who was involved, and how the data was downloaded); Ex. A

plaintiffs asked Magistrate Judge Wang to permit further discovery into OpenAI's reasons for deleting the datasets—reasons OpenAI has always maintained are privileged because the decision to delete "was made with a consultation with the company attorneys." Ex. B (Jan. 29, 2025 Trinh Dep. Tr.) 61:25–62:1.

Judge Wang acknowledged OpenAI's privilege assertions but nevertheless found that OpenAI had impliedly waived privilege. That holding is clearly erroneous and contrary to law; none of Judge Wang's three reasons for finding waiver survive scrutiny.

First, OpenAI has not waived privilege by partial disclosure, for it has never revealed any of its reasons for deleting the datasets. *See* Order at 15. Ever since the class plaintiffs first asked about OpenAI's reasons for deleting the datasets in January 2025, OpenAI has rebuffed those efforts by invoking attorney-client privilege. *See, e.g.*, Ex. B 61:19–23. Ignoring OpenAI's privilege assertions, Judge Wang pointed to a stray line in two off-topic and since-retracted discovery letters from early 2024 stating that the datasets were deleted "due to their non-use." Order at 16. Yet those remarks refer not to any of OpenAI's privileged reasons, but to the simple fact that OpenAI deleted the datasets after it stopped using them. In other words, non-use provided the occasion for OpenAI's discussions about deleting the datasets. Regardless, even if the letters had partially revealed one of OpenAI's reasons, Federal Rule of Evidence 502 would still foreclose waiver because OpenAI's inadvertent disclosure did not disadvantage the plaintiffs.

Second, OpenAI has not waived privilege by somehow changing its position. *See* Order at 16. OpenAI has consistently asserted that its *reasons* for deleting the datasets are privileged, even

58:1–60:18, 63:16–21 (discussing the lawyers and OpenAI employees involved in the decision to remove the datasets, the factual subject matter of those discussions, and when and where they took place; Ex. A 78:3–7, 86:13–88:13, 126:18–128:24 (discussing when the datasets were removed, by whom, and the technical steps taken for the removal); *see also* ECF 504 at 3 (noting that OpenAI "successfully located and produced" recovered copies of the datasets).

as OpenAI has accepted that the plaintiffs should have the opportunity to get discovery on the *fact* of non-use—an opportunity the plaintiffs have taken. Judge Wang accused OpenAI of inconsistency, suggesting that it had waffled on whether "non-use" was a privileged reason for deletion and whether the plaintiffs should get discovery on the subject. Order at 16. But Judge Wang's analysis downplayed the critical point that references to "non-use" in OpenAI's submissions refer to the non-privileged *fact* of non-use, not to a privileged *reason* for deleting the datasets. In any event, assuming OpenAI had changed its stance, that change is not the sort of "willful[l]" or "cavalier" violation of the federal or local rules that justifies forced waiver. *See In re Honeywell Int'l, Inc. Sec. Litig.*, 230 F.R.D. 293, 299 (S.D.N.Y. 2003) (quoting *AFP Imaging Corp. v. Philips Medizin Sys.*, 1993 WL 541194, at *3 (S.D.N.Y. Dec. 28, 1993)).

Third, OpenAI has not put privileged communications at issue simply by denying the plaintiffs' allegations of willfulness. *See* Order at 17, 19. Second Circuit precedent explicitly limits at-issue waivers to circumstances in which parties seek to "*rely* on privileged advice from … counsel to make [a] claim or defense." *In re County of Erie* (*Erie II*), 546 F.3d 222, 229 (2d Cir. 2008) (emphasis in original). The fact that "privileged information may be in some sense *relevant*" to a lawsuit is not enough. *Id.* (emphasis in original). And here, OpenAI has expressly disavowed any intent to "rely on any non-privileged reason for the removal of the Books1 and Books2 datasets in th[is] litigation." ECF 504 at 2. Judge Wang deemed that disclaimer irrelevant, finding an at-issue waiver based solely on OpenAI's denial that it acted willfully. But the "mere indication of a … defense" is never enough to waive privilege. *Erie II*, 546 F.3d at 229. Judge Wang's contrary rule threatens to "render … privilege a nullity" by giving "skillful pleader[s]" the ability to force privilege waivers with pleadings that bring "the adversary's knowledge and good faith into issue…." *Gardner v. Major Auto. Cos., Inc.*, 2014 WL 1330961, at *7 (E.D.N.Y. Mar.

31, 2014) (quoting *Chase Manhattan Bank N.A.* v. *Drysdale Secs. Corp.*, 587 F. Supp. 57, 58 (S.D.N.Y. 1984)).

This Court should set aside Magistrate Judge Wang's Order, hold that OpenAI's reasons for the removal of the datasets are privileged, that OpenAI has not waived that privilege, and end this case's time-consuming detour into facts that are irrelevant to the infringement claims at issue.

II. BACKGROUND

A. Class Plaintiffs Allege Copyright Infringement

OpenAI develops artificial-intelligence models, collectively referred to as GPT-N. OpenAI released its first model, "GPT-1," in 2018. Consolidated Class Action Compl. ("CCAC") ¶ 81, ECF 183. And, in November 2022, OpenAI released ChatGPT, a consumer-facing product built on OpenAI's models. CCAC ¶ 85. ChatGPT was initially powered by GPT-3.5 but was transitioned to another GPT model after a few months. CCAC ¶ 80; Ex. A 35:15–36:9. Class plaintiffs are authors and copyright holders of various books. CCAC ¶¶ 2, 166–292. The plaintiffs allege that OpenAI and Microsoft infringed copyrights in their books by downloading and copying their works to use in "training"—*i.e.*, developing— the GPT models. CCAC ¶¶ 2, 5–7, 13, 308–14. To support those allegations, the plaintiffs focus on the datasets OpenAI used to train its GPT models, which the plaintiffs assert contained unauthorized copies of their works. The class plaintiffs acknowledge that OpenAI has used different datasets over time. For instance, OpenAI has long used "Common Crawl," a "vast and growing corpus" of publicly available "web data." CCAC ¶¶ 111–12.

Beyond those sources, OpenAI briefly used two other datasets called "Books1" and "Books2" to train its GPT-3 and GPT-3.5 models. CCAC ¶¶ 111, 120. Former OpenAI researchers created those datasets using content downloaded from LibGen, a website that hosts digital copies of books. CCAC ¶¶ 113–16, 168, 182, 279; OpenAI's Answer to CCAC ("Answer")

¶ 119, ECF 754. The last model trained using Books1 and Books2, GPT-3.5, was trained in 2021 and discontinued by early 2023. *See* CCAC ¶¶ 84–88. In mid-2022, after OpenAI had already stopped using the datasets in training GPT models—and a full year before any lawsuit was filed challenging OpenAI's training on books—OpenAI deleted Books1 and Books2. *See* Ex. A 78:3–5; ECF 504.

Despite the small role Books1 and Books2 played in developing two discontinued GPT models, the class complaint devotes considerable attention to those datasets. Class plaintiffs claim that LibGen hosted "unauthorized copies of copyright protected content," including their books. CCAC ¶¶ 113–16, 168, 182, 279. And because former OpenAI researchers used LibGen to create Books1 and Books2, class plaintiffs assume that their works also formed part of those datasets. *See, e.g.*, CCAC ¶¶ 113–16. Class plaintiffs then make the further leap of asserting that OpenAI "knew that [its] 'training' data included texts protected by copyright" and so "willfully" infringed those copyrights when it supposedly used the class plaintiffs' works without permission. CCAC ¶ 121.

As that summary illustrates, the class plaintiffs' allegations focus on OpenAI's purported *use* of unauthorized copies of the plaintiffs' works to train certain GPT models. Notably absent from the Class Complaint are any allegations about OpenAI's later decision to *delete* the Books1 and Books2 datasets. Class plaintiffs have not alleged that the deletion itself somehow constituted copyright infringement or gives rise to some other cause of action. *See generally* CCAC.

OpenAI denies that it infringed the plaintiffs' works, whether willfully or otherwise. *See generally* Answer. OpenAI admits that "the books1 and books2 datasets included content from LibGen" and that those "datasets … were used in training …." Answer ¶ 119. But OpenAI lacks any "knowledge or information" about whether the class plaintiffs' works were hosted on LibGen.

E.g., Answer ¶¶ 182–83. And while OpenAI denies that it acted willfully, it did not assert good faith or advice of counsel among its list of 21 affirmative defenses. *See* Answer at ¶¶ 121, 175, 200, 211, 221, 226, 233, 242, 247, 253, 266, 276, 286, 301, 312 (denying willfulness allegations); pp. 39–42 (listing affirmative defenses). As a result, neither good faith nor advice of counsel are at issue in this case. Like the Class Complaint, OpenAI's Answer made no mention of the deletion of the Books1 and Books2 datasets.

B. OpenAI Asserts Attorney-Client Privilege Over Its Reasons for Deleting the Datasets.

The present discovery dispute concerns OpenAI's assertion that attorney-client privilege protects its reasons for deleting the Books1 and Books2 datasets in 2022. Though class plaintiffs have had ample opportunity to discover non-privileged facts about OpenAI's *acquisition* and *use* of the datasets, the class plaintiffs have nevertheless pressed for additional discovery into OpenAI's privileged reasons for *deleting* the datasets years later—an issue not remotely relevant to plaintiffs' infringement theory or any other claim in the Class Complaint. *See* ECF 504 at 1–2.

OpenAI has consistently rebuffed those efforts by asserting privilege. The class plaintiffs first inquired into the reasons why OpenAI deleted the datasets at the January 29, 2025 30(b)(6) deposition of Associate General Counsel Michael Trinh. When class plaintiffs' counsel asked Mr. Trinh whether he knew "why Books1 and Books2 were deleted," OpenAI's counsel objected that the question "call[ed] for privileged information." Ex. B (Jan. 29, 2025 Trinh Dep. Tr.) 61:19–23. Mr. Trinh confirmed that the decision to delete the datasets "was made with a consultation with the company attorneys," so he was unable to "answer further without revealing privileged information." Ex. B 61:25–62:3. That assertion of privilege could not have been clearer.

Since that deposition, OpenAI has consistently asserted privilege over its reasons for deleting the Books1 and Books2 datasets. For example, in a May 20, 2025 joint status update,

OpenAI argued that "the reasons for the deletion [of Books1 and Books2] are privileged." ECF 53-3 at 9. And at a hearing on privilege a few days later, on May 27, OpenAI reiterated its privilege assertion even as it indicated that it did not object to discovery on the *fact* that the datasets were not in use when OpenAI deleted them. ECF 109 at 70:4–12 ("[J]ust because there may be an element of this issue that is not privileged, doesn't mean that every aspect of the issue is not privileged.").

On June 13, 2025, OpenAI submitted revised versions of two 2024 discovery letters to clarify that those letters did not reveal any of OpenAI's privileged reasons for deleting the datasets. OpenAI sent the original letters on March 22 and April 16, 2024, to inform opposing counsel and the Court respectively that OpenAI was attempting to locate and produce copies of the Books1 and Books2 datasets. ECF 188-2; 188-4. The letters explained that years before this litigation began OpenAI had deleted the unused datasets, rendering production difficult. ECF 188-2; 188-4. The letters also mentioned in passing that OpenAI deleted the datasets "due to their non-use," ECF 188-2; 188-4—a phrase meant to describe the simple fact that OpenAI deleted the datasets after it stopped using them, ECF 109 at 70:4–6. After the class plaintiffs seized on the letters the following year as purported evidence of OpenAI's reasons for deleting the datasets, OpenAI revised the letters to omit the phrase "due to their non-use," thus avoiding any ambiguity about whether the letters had disclosed privileged information. ECF 188.

Throughout June and July 2025, OpenAI continued to assert privilege. On June 29, OpenAI raised a privilege objection to class plaintiffs' new 30(b)(6) deposition notice, which called for "reasons why each version of Books1 and/or Books2 was deleted, including 'non-use.'" ECF 413-10 at 8. On July 25, at Mr. Trinh's next 30(b)(6) deposition, he again asserted privilege over the reasons for deletion and explained that in-house counsel "led" the decision to remove

Books1 and Books2 in summer 2022. Ex. A 59:25–60:18. And on July 30, OpenAI represented in an email that it "does not believe that there are non-privileged reasons for the deletion" of the datasets. ECF 413-11 at 2.

C. Magistrate Judge Wang Orders OpenAI to Disclose *All* Communications Relating to Reasons for Deletion

On July 30, 2025, the class plaintiffs filed a letter motion to compel OpenAI to produce "materials" related to its reasons for deleting the Books1 and Book2 datasets. ECF 413 at 1. After initial responses from OpenAI, ECF 428; ECF 504, and a supplemental round of briefing from both parties, ECF 615; ECF 616, on November 24, Magistrate Judge Wang issued an Opinion and Order granting the plaintiffs' motion to compel "with respect to privilege waiver," Order at 27.[2]

Judge Wang assumed for purposes of the analysis that OpenAI's reasons for deleting the datasets are protected by attorney-client privilege. Order at 16. And Judge Wang did not identify any privileged communication that OpenAI voluntarily disclosed that would give rise to waiver. Judge Wang nevertheless held that OpenAI had waived privilege in three different ways: (1) by "disclosing a 'privileged reason' for the deletion" when it said that the datasets were "deleted due to their non-use" in the 2024 discovery letters, Order at 15 (capitalization altered); (2) by "making a moving target of its privilege assertions" when it offered purportedly different positions on whether "non-use" was a privileged reason for deleting the datasets, Order at 16; and (3) by putting its state of mind "at issue" in this case when it denied the class plaintiffs' allegations that OpenAI acted "willfully," Order at 17, 19–23.

Accordingly, Magistrate Judge Wang found that OpenAI had waived privilege over "all communications in 2022 related to the reasons for the deletion of (1) the Books1 and Books2

[2] Magistrate Judge Wang rejected the plaintiffs' argument that the crime-fraud exception to attorney-client privilege applies. Order at 24–27.

datasets and (2) all internal references to LibGen" and ordered the production of those communications by December 8, 2025. Order at 24, 27-28. Magistrate Judge Wang further ordered OpenAI to identify all in-house lawyers who participated in discussions regarding deletion of Books1 and Books2 and, by December 19, 2025, to make those lawyers available for two-hour depositions. Order at 28.

On December 4, 2025, Magistrate Judge Wang stayed the production of documents pending resolution of OpenAI's Rule 72 objections, and extended to January 30, 2026, the date by which in-house counsel must be available for depositions. ECF 910 at 5–6.

III. STANDARD OF REVIEW

Where a magistrate judge decides a "matter not dispositive of a party's claim or defense," Fed. R. Civ. P. 72(a), a "district court must 'modify or set aside any part of a magistrate judge's order that is clearly erroneous or is contrary to law.'" *Sekisui Am. Corp. v. Hart*, 945 F. Supp. 2d 494, 502 (S.D.N.Y. 2013) (brackets omitted) (quoting Fed. R. Civ. P. 72(a)). An order is contrary to law if it "fails to apply or misapplies relevant statutes, case law, or rules of procedure …." *Id.* (citation omitted). "Findings are clearly erroneous when the reviewing court is firmly convinced the lower court decided an issue in error." *Catskill Dev., L.L.C. v. Park Place Ent. Corp.*, 206 F.R.D. 78, 86 (S.D.N.Y. 2002).

IV. ARGUMENT

Magistrate Judge Wang's Order should be set aside. Both OpenAI's consistent privilege assertions and controlling law sharply limiting implied privilege waivers compel reversal.

A. OpenAI's Reasons for Deleting the Datasets Are Privileged

Attorney-client privilege is among our legal system's oldest and most important evidentiary rules. The privilege ensures "full and frank communication between attorneys and their clients and thereby promote[s] broader public interests in the observance of law and

administration of justice." *Upjohn Co. v. United States*, 449 U.S. 383, 389 (1981). To that end, the privilege shields all "communications" "between a client and his or her attorney" that "are intended to be, and in fact were, kept confidential" and were made "for the purpose of obtaining or providing legal advice." *United States v. Mejia*, 655 F.3d 126, 132 (2d Cir. 2011).

Under that standard, attorney-client privilege plainly protects OpenAI's reasons for deleting the Books1 and Books2 datasets. As Mr. Trinh testified at his depositions, "discussion[s]" about deleting Books1 and Books2 were "led by legal," Ex. A 60:7–10, and ultimately the decision to delete "was made with a consultation with the company attorneys," Ex. B 61:25–62:3. Because OpenAI's reasons for deleting the documents are inextricably bound up with legal "recommendations," attorney-client privilege protects them. *See In re County of Erie*, 473 F.3d 413, 417 (2d Cir. 2007) (*Erie I*). Disclosing those reasons would necessarily reveal confidential communications between OpenAI and its lawyers.

Magistrate Judge Wang did not hold otherwise. Though Judge Wang initially suggested it was "unclear" whether OpenAI's reasons for deletion might be non-privileged facts, Order 15–16, Judge Wang did not definitively resolve that issue. Instead, Judge Wang's analysis proceeded on the assumption that OpenAI's "reasons" for deleting the datasets are privileged. Order 16–23. Because Judge Wang did not definitively resolve whether OpenAI's reasons are attorney-client privileged, this Court must address the question "de novo." *See MacNamara v. City of New York*, 249 F.R.D. 70, 82 & n.7 (S.D.N.Y. 2008).

OpenAI's reasons for deleting the datasets were not mere "underlying facts" that OpenAI happened to disclose to counsel. *See Upjohn Co.*, 449 U.S. at 395–96 ("privilege … does not protect disclosure of the underlying facts"). OpenAI's reasons are instead the *product* of "consultation with the company attorneys." ECF 413-8 at 18–19. And a client's reasons for acting

based on legal advice are privileged. *See, e.g.*, *Sanofi-Synthelabo v. Apotex Inc.*, 299 F. Supp. 2d 303, 307–08 (S.D.N.Y. 2004) (The "communications" between a company and its counsel about "the reasoning behind" a decision to amend a patent were privileged communications because they were "clearly related" to the provision of "legal services."), *on reconsideration*, 363 F. Supp. 2d 592 (S.D.N.Y. 2005). So too here, OpenAI's reasoning is part and parcel of privileged legal advice.

B. OpenAI Did Not Waive Privilege

Even though OpenAI's reasons for deletion are privileged, Magistrate Judge Wang nevertheless found that OpenAI had implicitly "waived" that privilege. Order at 16. To support that conclusion, Judge Wang advanced not one but three different theories. OpenAI waived privilege, Judge Wang reasoned, either by "disclosing a 'privileged reason' for the deletion," Order at 15, by "making a moving target of its privilege assertions," Order at 16, or by putting privileged communications "at issue" in this case, Order at 17. None of those reasons withstands scrutiny.

1. OpenAI Did Not Waive Privilege by Partial Disclosure

Despite the name, implied privilege waivers often involve "no representation, express or implied, that [a party] intends to surrender its privilege." *John Doe Co. v. United States*, 350 F.3d 299, 302 (2d Cir. 2003). Implied waivers thus "possess the potential to weaken attorney-client trust" and courts should approach them with "caution." *Erie II*, 546 F.3d at 228.

Deploying that cautious approach, courts have found that a party may waive privilege over the entirety of a communication by selectively disclosing misleading portions of it. *In re Grand Jury Proc.*, 219 F.3d 175, 182 (2d Cir. 2000). Here, by contrast, OpenAI disclosed *no* privileged information at all, so the partial-disclosure rule does not apply. And even if OpenAI had disclosed privileged information, Federal Rule of Evidence 502 would foreclose waiver because any disclosure was inadvertent and caused no unfairness to the class plaintiffs.

a. The record demonstrates that OpenAI has never divulged any of its reasons for deleting the Books1 and Books2 datasets. Throughout this litigation, OpenAI has consistently maintained that all its reasons for deleting the datasets are privileged. *See* ECF 504 at 2. When plaintiffs for the first time sought information about those reasons at Mr. Trinh's January 29, 2025 deposition, Mr. Trinh replied that he could not "answer further without revealing privileged information" because the decision to delete Books1 and Books2 "was made in consultation with the company attorneys." Ex. B 61:25–62:1. OpenAI has maintained that stance ever since. *See, e.g.*, ECF 53-3 at 9 (joint status report reflecting OpenAI's position that the "reasons for the deletion are privileged"); Ex. A 60:7–10 ("the discussion to delete was led by legal").

Minimizing those consistent assertions, Judge Wang instead homed in on a single clause—"[the] datasets were deleted due to their non-use"—repeated in two off-topic discovery letters that OpenAI sent in 2024. Judge Wang found that these statements partially disclosed one of OpenAI's reasons for deleting the datasets. Order at 15–17. And in Judge Wang's view, that partial disclosure justified a waiver of OpenAI's privilege over *all* communications concerning its reasons for deleting the datasets. Order at 17.

Judge Wang reached that conclusion only by ignoring critical context. The letters were not intended to address OpenAI's reasons for deleting the datasets. Instead, the letters provided opposing counsel and the Court with discovery updates on OpenAI's efforts to locate the datasets. The letters thus explained that OpenAI could not readily produce Books1 and Books2 because OpenAI had deleted them in 2022 after it stopped using them and detailed OpenAI's best efforts to "locate any additional copies of the datasets or other documents from which the datasets can be reconstructed," ECF 188-1 at 4; ECF 188-3 at 3—efforts that were ultimately successful.

In context, the phrase "due to their non-use" did not disclose OpenAI's sensitive and confidential deliberations over deleting the datasets. The phrase should be read as OpenAI intended it: to convey the fact that OpenAI deleted two datasets that had fallen out of use. "[D]ue to" was a natural (if in-artful) way to express the point that non-use formed part of the factual backdrop OpenAI confronted when deciding whether to remove the datasets. In that limited sense, "non-use" was a "cause of the deletions," as OpenAI's counsel later recognized. ECF 109 at 70:4–6. OpenAI confirmed that the letters did not disclose a privileged reason for deletion by submitting revised versions of the letters omitting the "due to" language that the plaintiffs and Judge Wang have seized on. ECF 188.

Judge Wang addressed the critical distinction between the fact of non-use and OpenAI's reasons for deletion only in a footnote, suggesting that the "ordinary meaning of the words 'due to'"—*i.e.*, "[b]ecause of"—"belied" the notion that "'due to non-use' was not intended to describe a 'reason' for the deletion of Books1 and Books2." Order at 15 n.14 (citing Due To, *The Am. Heritage Coll. Dictionary* (5th ed. 2022)). But that dictionary definition is beside the point. Of course, in a loose sense, OpenAI deleted the datasets "because" they were no longer being used. That fact formed a critical part of the backdrop behind OpenAI's decision to delete the datasets. The non-use of Books1 and Books2 provided the occasion for discussions with in-house counsel about whether OpenAI should remove the datasets. But that does not mean that "non-use" was among OpenAI's privileged *reasons* for deletion after discussing the issue with counsel. In all events, "words must be read and interpreted in their context, not in isolation." *Sw. Airlines Co. v. Saxon*, 596 U.S. 450, 455 (2022) (quotation marks omitted). And in context, the 2024 discovery letters do not disclose a privileged reason for deleting the datasets. *Supra*, pp. __.

Judge Wang's footnote also suggested—without record citation—that OpenAI must view "non-use" as privileged because it had "refus[ed] to testify about 'non-use' … based on its potential to waive privilege" and used "privilege to block any inquiry into 'non-use….'" Order at 15 n. 14. Judge Wang appears to be referring to Mr. Trinh's second deposition, where counsel stopped Mr. Trinh from answering questions about nonprivileged facts regarding the reasons OpenAI deleted the datasets. ECF 413-1 at 107:5–14. But counsel only gave that advice because the plaintiffs would not agree in advance that disclosure of nonprivileged facts would not constitute a waiver. ECF 413-1 at 107:15–18. OpenAI's careful protection of its privilege does not prove that it regards the underlying facts, including non-use, as also privileged. Nor does OpenAI's cautious approach to privilege prove that it waived privilege. Regardless, the plaintiffs have had and taken an opportunity to get discovery on non-use. *See, e.g.*, Ex. A 62:7–63:15 ("Q: And you're not aware of the datasets being used for anything else at the time of … oral discussions [about the deletion of the datasets]? A: Correct.").

Because OpenAI disclosed zero privileged information, waiver based on partial disclosure is unwarranted. The partial-waiver rule applies only to incomplete disclosures of "*privileged* communications," and thus has no role to play where, as here, a party discloses no privileged information at all. *See Erie II*, 546 F.3d at 228 (emphasis added). "'At issue' waiver is concerned with the selective disclosure of privileged material, not the disclosure of non-privileged material." *Aiossa v. Bank of Am.*, 2011 WL 4026902, at *5 (E.D.N.Y. Sept. 12, 2011); *accord Murtha v. New York State Gaming Comm'n*, 2020 WL 5504311, at *2 (S.D.N.Y. Sept. 9, 2020); *Maiurano v. Cantor Fitzgerald Sec.*, 2021 WL 4993060, at *2 n.1 (S.D.N.Y. Oct. 27, 2021). Courts across the country have thus recognized that "the disclosure of non-privileged material never waives the attorney-client privilege." *In re FirstEnergy Corp.*, 154 F.4th 431, 443 (6th Cir. 2025); *accord*

E.E.O.C. v. George Washington Univ., 502 F. Supp. 3d 62, 87 (D.D.C. 2020); *Kovacs v. Hershey Co.*, 2006 WL 3054167, at *3 (D. Colo. Oct. 25, 2006). This court should apply the same well-established rule here and find no waiver based on OpenAI's disclosure of the non-privileged fact of non-use.

For the same reason, the plaintiffs are wrong to suggest that OpenAI waived privilege by disclosing *non-privileged* communications pursuant to Magistrate Judge Wang's October 1, 2025 Order. See ECF 907 at 7; ECF 618. As Judge Wang held, the communications produced under that Order are not privileged, ECF 618, so disclosing them could not have provoked a waiver over other privileged communications, *see In re FirstEnergy Corp*., 154 F.4th at 443. In any case, none of the disclosed communications reveal OpenAI's reasons for deleting the datasets. For instance, an OpenAI employee's recognition that "legal reasons" support deletion plainly does not reveal what those reasons are and instead only underscores that OpenAI's reasons are protected by attorney-client privilege. ECF 907-2.

b. Even if Magistrate Judge Wang were correct on the record, Federal Rule of Evidence 502 would still foreclose waiver, for two reasons.

First, any disclosure was inadvertent. Under Rule 502(b), "inadvertent disclosure of protected communications or information … does not constitute a waiver if the holder took reasonable steps to prevent disclosure and also promptly took reasonable steps to rectify the error." Fed. R. Evid. 502 note. That describes this case to a T. When OpenAI represented that it deleted the datasets "due to their non-use," it did not intend to disclose any privileged information. *See* ECF 504. OpenAI also took "reasonable steps to prevent disclosure," Fed. R. Evid. 502(b)(2), as illustrated by the fact that in all the documents produced and depositions held in this case, snippets from two off-topic letters are the *only* disclosures even arguably related to OpenAI's reasons for

deleting the datasets. *See* ECF 188. OpenAI then made a "reasonable" effort "to rectify the error," Fed. R. Evid. 502(b)(3), by submitting revised versions of the letters as soon as OpenAI realized that the plaintiffs misread them as disclosing a privileged reason for deletion. ECF 188; *see Briese Lichttechnik Vertriebs GmbH v. Langton*, 272 F.R.D. 369, 372–73 (S.D.N.Y. 2011) (no waiver under Rule 502(b) where a party later sought to retrieve inadvertently disclosed, privileged information).

Second, OpenAI has not sought to use privilege as a sword to disadvantage the plaintiffs. Under Rule 502(a), even *intentional* disclosures do not waive privilege over undisclosed communications concerning the same "subject matter" unless "fairness requires a further disclosure … to prevent a selective and misleading presentation of evidence to the disadvantage of the adversary." Fed. R. Evid. 502(a) note; *see also In re Grand Jury Proc.*, 219 F.3d at 182 ("[I]mplied waiver may be found where" a party "partially disclose[s] privileged communications" that "*in fairness* requir[e] examination of [other] protected communications." (emphasis in original) (quoting *United States v. Bilzerian*, 926 F.2d 1285, 1292 (2d Cir. 1991)). But there is no risk of unfairness here. OpenAI has expressly disclaimed any intent to rely on evidence about its reasons for deleting the datasets, defusing any risk of an incomplete or misleading presentation of the evidence. Nor, contrary to the plaintiffs' suggestion, has OpenAI "selectively permitt[ed] fact witnesses" to testify about OpenAI's reasons for deleting the datasets. ECF 907 at 8. Tellingly, the plaintiffs' best evidence of that purported tactic—the deposition of OpenAI employee Long Ouyang—involves no partial disclosure of any reason for deletion. Instead, Mr. Ouyang testified that he *assumed*, without any personal knowledge, that OpenAI deleted the datasets for "copyright" reasons after in-house counsel instructed him to delete the datasets *without* specifying any reason

why. ECF 907-1 (the "entire contents" of the message from counsel was "an instruction to delete.").

Where, as here, a partial disclosure "has not afforded … any tactical advantage in [the] litigation," and there is no potential for evidence to "be used selectively at trial," courts routinely find that Rule 502(a) does not permit waiver as to undisclosed privileged communications. *Swift Spindrift, Ltd. v. Alvada Ins., Inc*., 2013 WL 3815970, at *5 (S.D.N.Y. July 24, 2013); *see, e.g.*, *In re Keurig Green Mountain Single-Serve Coffee Antitrust Litig*., 2020 WL 3034792, at *5 (S.D.N.Y. June 5, 2020). This Court should follow that well-trodden path here.

2. OpenAI Did Not Waive Privilege by Changing Positions

Judge Wang also concluded that "OpenAI has waived privilege by making a moving target of its privilege assertions." Order at 16. Once again, neither the record nor the law supports that conclusion. The record shows that OpenAI has consistently asserted privilege. And the law holds that privilege waivers based on litigation conduct are reserved for "willful[l]" or "cavalier" violations of the federal or local rules—not subtle changes in a party's position. *See In re Honeywell Int'l*, 230 F.R.D. at 299 (citation omitted).

a. The record is clear: OpenAI has always maintained that its *reasons* for deleting the datasets are privileged, but that the *fact* of non-use is not. Thus, OpenAI has told the Court that "the reasons for the deletion [of the datasets] are privileged," ECF 53-3 at 9, and objected to discovery on "all reasons why each version [of the datasets] was deleted," ECF 413-10 at 8, even as OpenAI has "not assert[ed] privilege … [or] blocked plaintiffs from exploring the question of nonuse," ECF 109 at 70:4–6.

Ignoring that straightforward interpretation of the record, Judge Wang instead decided that OpenAI had changed its position twice: "OpenAI initially [in its 2024 discovery letters] stated … that Books1 and Books2 were deleted 'due to their non-use,' … then [at a January 2025 deposition]

asserted that all the 'reasons' for the deletion of Books1 and Books2 were privileged … then [at a May 27, 2025, hearing] claimed that not all aspects of the 'reasons' for deletion were privileged, including the question of non-use…." Order at 16.

But again, that mischaracterization overlooks the critical distinction between the *fact* of non-use and OpenAI's privileged *reasons* for deletion. As discussed, the 2024 discovery letters used the phrase "due to … non-use" merely to indicate the fact that OpenAI was not using the datasets when it deleted them. *See supra*, pp. __. In reality, OpenAI had no occasion to assert privilege until the plaintiffs first inquired about OpenAI's reasons for deleting the datasets at Mr. Trinh's January 2025 deposition the following year. *See* Ex. B 61:19–62:3. From that moment on, OpenAI has asserted that its reasons are privileged. At the same time, OpenAI has acknowledged that the plaintiffs are free to inquire into the fact of non-use. As OpenAI explained at the May 27, 2025, hearing, while counsel asserted privilege over OpenAI's reasons, OpenAI has "not blocked plaintiffs from exploring the question of nonuse of the data set" because that factual "element of th[e] issue … is not privileged." ECF 109 at 70:4–12. That was and remains true.

b. Even if OpenAI has changed its position on privilege, any shift is miles away from the kind of litigation misconduct that merits a forced privilege waiver. Although "willful[l]" or "cavalier" violations of the federal or local discovery rules may waive privilege, small clarifications in parties' position do not. *In re Honeywell Int'l*, 230 F.R.D. at 299.[3] OpenAI has

[3] *Accord In re DG Acquisition Corp.*, 151 F.3d 75, 84 (2d Cir. 1998) (finding no privilege waiver despite a failure to assert privilege within the time required by the Federal Rules of Civil Procedure where the party did "not flagrantly ignore[] communications and court orders" or otherwise show "willfu[l] or cavalier disregard" for the Rules" (citation omitted)); *RMED Int'l, Inc. v. Sloan's Supermarkets, Inc.*, 2003 WL 41996, at *3 (S.D.N.Y. Jan. 6, 2003) (explaining that courts "must … examine, the nature of the violation, its willfulness or cavalier disregard for the rule's requirements, and the harm which results to other parties" (citation omitted)); 24 Wright & Miller,

not, for instance, "fail[ed] to serve indices of privileged documents in a timely and proper manner" as required by the Federal Rules. *S.E.C. v. Yorkville Advisors, LLC*, 300 F.R.D. 152, 167 (S.D.N.Y. 2014). Nor has OpenAI sought to "re-engineer privilege logs to align [its] privilege assertions with [its] legal arguments." *In re Honeywell Int'l*, 230 F.R.D. at 299. Even accepting Judge Wang's interpretation of the facts for the sake of argument, OpenAI at most imprecisely articulated the metes and bounds of one privilege assertion over one specific topic.

The only case Judge Wang cited to support waiver in these circumstances actually cuts in OpenAI's favor. *See* Order at 14 (citing *Chelsea Hotel Owner LLC v. City of New York*, 2024 WL 323216, at *2 (S.D.N.Y. Jan. 29, 2024), *aff'd*, 2025 WL 880530 (S.D.N.Y. Mar. 21, 2025)). In *Chelsea Hotel*, Judge Lehrburger found that a party had *not* waived privilege by making "fluid" assertions of privilege "with certain privilege claims added when they had not previously been included and others removed when they previously had been included." 2024 WL 323216, at *2. If waiver was not justified there, it certainly is not justified here. Unlike the defendant in *Chelsea Hotel*, OpenAI has not added or subtracted privilege claims. The only purported "fluidity" is over the *scope* of OpenAI's privilege assertion as to its reasons for deleting the datasets. So the plaintiffs have always been on notice of the specific privilege OpenAI is asserting and what topic the privilege concerns.

The plaintiffs likewise fail to identify any authority supporting waiver based on OpenAI's conduct in this case. ECF 907 at 5. The best the plaintiffs can muster is a solitary and easily distinguishable out-of-circuit decision refusing to stay an order to compel. *See Digital-Vending Servs. Int'l, LLC v. Univ. of Phoenix Inc.*, 2010 WL 11450510 (E.D. Va. Apr. 22, 2010). There,

---

Fed. Prac. & Proc. Evid. § 5507.2 (1st ed. 2025 update) (noting that for even "calculated and purposeful" rules violations, "[t]he correct action is to sanction the party, rather than find a waiver" (citation omitted)).

the court held that a party had likely waived privilege through a "pattern of delay" and efforts to obscure "genuine issues related to discovery"—including "at the eleventh hour, add[ing] hundreds of new privilege claims that were never before asserted." *Id.* at *1 & n.2, *3. OpenAI has done nothing like that here.

3. OpenAI Did Not Waive Privilege by Putting Privileged Communications at Issue

Magistrate Judge Wang buttressed her initial two waiver theories with a third, finding that OpenAI waived privilege merely by denying allegations that it acted willfully. Order at 17–18. That holding is contrary to controlling precedent and threatens to eviscerate privilege in any case in which a defendant's state of mind is at issue.

a. Second Circuit precedent limits at-issue waivers to circumstances in which parties seek to "*rely* on privileged advice from … counsel to make [a] claim or defense." *Erie II*, 546 F.3d at 229 (emphasis in original). The fact that "privileged information may be in some sense *relevant*" to a lawsuit is not enough, nor is a "mere indication of a claim or defense." *Id.* (emphasis in original). Accordingly, courts will find at-issue waivers only if privileged communications are "indispensable to a party's claims or defenses," as where "the very subject of privileged communications is critically relevant to the issue to be litigated" and "there is no other source of direct proof …." *Chin v. Rogoff & Co.*, 2008 WL 2073934, at *5 (S.D.N.Y. May 8, 2008) (citation omitted).

The facts here come nowhere close to meeting that standard. OpenAI has not and will not seek to "rely" on privileged advice to make a "claim or defense" in this case. *Erie II*, 546 F.3d at 229 (emphasis omitted). Quite the contrary, OpenAI has repeatedly told the plaintiffs and the Court that "OpenAI *does not* intend to introduce evidence of the advice of counsel either affirmatively or in rebuttal to the … [plaintiffs'] claim[s]," ECF 413-11 at 2 (emphasis added), nor

will OpenAI "rely on any non-privileged reason for the removal of the Books1 and Books2 datasets in th[is] litigation," ECF 504 at 4.

Despite those disclaimers, Judge Wang found that OpenAI had "waived privilege over all communications regarding the 'reasons' for the deletion of Books1 and Books2 by putting its good faith and state of mind at issue." Order at 17 (emphasis omitted). Yes, Judge Wang acknowledged, OpenAI had "removed its good faith affirmative defense" in its Answer to the Class Complaint. Order at 18–19. Even so, Judge Wang reasoned, OpenAI necessarily "argue[d] that it acted in good faith" by denying the plaintiffs' allegations that its infringement was willful. Order at 19. And, in Judge Wang's view, that denial waived privilege because attorney-client "communications regarding the reasons for deleting Books1 and Books2 could be probative of OpenAI's willfulness." Order 19–20.

That reasoning badly muddles copyright law and drastically expands at-issue privilege waivers. Copyright law leaves no doubt that OpenAI has not asserted a good-faith defense by denying allegations of willfulness. Though willfulness and good faith both go to the amount of damages for infringement, they are distinct inquiries with distinct burdens. By statute, if a court finds that an "infringement was committed willfully" it may award increased statutory damages up to $150,000 per infringement. 17 U.S.C. § 504(c)(2). Conversely, if a court finds that the infringer acted in good faith—*i.e.*, "was not aware and had no reason to believe that his or her acts constituted an infringement of copyright"—then the court may limit damages to "$200" per infringement. *Id.* But those inquiries are not coterminous. "'Innocent' intent … is more than just the absence of willfulness": Whereas willfulness requires the *plaintiff* to prove that a defendant acted "with knowledge that [his] conduct constitutes copyright infringement," innocence requires the *defendant* to prove that he acted in good faith. *Nat'l Football League v. PrimeTime 24 Joint*

Venture, 131 F. Supp. 2d 458, 476 (S.D.N.Y. 2001) (citation omitted); *accord Fitzgerald Publ'g Co. v. Baylor Publ'g Co.*, 807 F.2d 1110, 1115 (2d Cir. 1986).

Accordingly, copyright defendants may deny a plaintiff's allegation of willfulness while simultaneously declining to assume the burden of proving good faith as an affirmative defense. That is what happened here: OpenAI's Answer denied the plaintiff's charge that OpenAI acted with "knowledge" or "reckles[s] disregar[d]," *Bryant v. Media Right Prods., Inc.*, 603 F.3d 135, 143 (2d Cir. 2010) (quoting *Twin Peaks Prods., Inc. v. Publ'ns Int'l, Ltd.*, 996 F.2d 1366, 1382 (2d Cir. 1993))—but the Answer did not assert an affirmative good-faith defense, as Judge Wang recognized, Order at 19.

Applying the correct legal standard, it is clear that OpenAI did not waive privilege by denying willfulness allegations. Courts in this district have long recognized that a plaintiff's presentation of a "legal claim implicating the [defendant's] knowledge" is not "sufficient to trigger a broad waiver of the [defendant's] attorney client privilege." *Green v. Beer*, 2010 WL 3422723, at *6 (S.D.N.Y. Aug. 24, 2010) (quoting *Laborers Local 17 Health Benefit Fund v. Philip Morris, Inc.*, 1998 WL 414933, at *2 (S.D.N.Y. July 23, 1998)). Instead, even where a defendant is "alleged to have acted willfully," the "advice of counsel is placed in issue where the client asserts a claim or defense, and attempts to prove that claim or defense by disclosing or describing an attorney client communication." *Rhone-Poulenc Rorer Inc. v. Home Indem. Co.*, 32 F.3d 851, 863 (3d Cir. 1994), *cited favorably by Erie II*, 546 F.3d at 227. OpenAI's denial of the allegation that acted knowingly is plainly insufficient to waive privilege.

Finding implied waiver based only on a defendant's denial of willfulness would render attorney-client privilege "a nullity" in almost any case in which the plaintiff's pleadings implicate the defendant's state of mind. *Standard Chartered Bank PLC v. Ayala Int'l Holdings (U.S.) Inc.*,

111 F.R.D. 76, 84–85 (S.D.N.Y.1986). If a defendant can put privileged communications at issue merely by denying allegations of a guilty mind—without any reference to privileged communications or the attorney-client relationship—then clever plaintiffs can always put defendants to an impossible choice: decline to contest the allegation or waive privilege over *any* related attorney-client communications. Precisely to avoid giving "skillful pleader[s] the ability to render … privilege a nullity," courts in this circuit have long refused to permit a party to "force its adversary to waive its privilege implicitly with a pleading that brought the adversary's knowledge and good faith into issue…." *Gardner*, 2014 WL 1330961, at *7 (quoting *Chase Manhattan Bank N.A.*, 587 F. Supp. at 58).

b. Even assuming OpenAI had asserted a good-faith defense, as Judge Wang suggested, Order at 21–22, the "mere indication" of a defense is not enough to waive privilege. *Erie II*, 546 F.3d at 229.

To be sure, "the assertion of a good-faith defense" may sometimes "cal[l] forth the possibility of implied waiver." *Id.* at 228–29 (citing *Bilzerian*, 926 F.2d at 1292). But to avoid open[ing] a great number of privileged communications to claims of at-issue waiver," the Second Circuit has clarified that invoking a defense, without more, is "insufficient to place legal advice at issue." *Id.* at 229. Instead, a good-faith defense waives privilege only where the defendant had "the advice of counsel on *the very issue* on which [he] asserts good faith …." *Scott v. Chipotle Mexican Grill, Inc*., 67 F. Supp. 3d 607, 618 (S.D.N.Y. 2014) (emphasis added). That is the case where, for instance, a defendant argues that he made a good-faith attempt to comply with the securities laws but seeks to shield communications with his attorney about the very funds at issue. *Bilzerian*, 926 F.2d at 1292; *accord Okla. Firefighters Pension & Ret. Sys. v. Musk*, 2025 WL 2982003, at *4, *6 (S.D.N.Y. Oct. 23, 2025) (finding waiver based on defendants' "inten[t] to

assert good faith belief in the lawfulness of their actions … after they first consulted with counsel" on the "very conduct" at issue).

Those ground rules foreclose waiver here because OpenAI's reasons for deleting the datasets do not concern the "very issue" in dispute. Any good-faith or willfulness inquiry in this case will focus on OpenAI's state of mind regarding the alleged infringing conduct—*i.e.*, OpenAI's purported *use* of copyrighted material in training GPT-N. That is why the Class Complaint focuses on what OpenAI knew or should have known when it *created* the Books1 and Books2 datasets. *See, e.g.*, CCAC ¶¶ 111–16, 120–21. By contrast, OpenAI's state of mind when it *deleted* the datasets years later is simply not at issue—a point underscored by the Class Complaint's failure to even mention the deletion of Books1 or Books2. Even if it "might be useful" to the defendants to have access to OpenAI's privileged communications, those communications are not "critical" to the claims in this case and thus do not fall within the at-issue waiver rule. *Leviton Mfg. Co. v. Greenberg Traurig LLP*, 2010 WL 4983183, at *5 (S.D.N.Y. Dec. 6, 2010), *objections overruled*, 2011 WL 2946380 (S.D.N.Y. July 14, 2011).

In reaching a contrary conclusion, the Order placed far too much weight on a single distinguishable case. *See* Order at 20–23 (citing *Arista Records, LLC v. Lime Grp., LLC*, 2011 WL 1642434, at *2 (S.D.N.Y. Apr. 20, 2011)). *Arista Records* concerned a copyright-infringement suit against LimeWire. After the court found infringement, the defendants argued that they should pay reduced damages because they "had a good faith belief with respect to … the lawfulness of their conduct in operating LimeWire." 2011 WL 1642434, at *1. In support of that defense, one of the defendants, Mark Gorton, testified to his belief that "LimeWire was not at great legal risk" at the time of the infringement. *Id.* But "whenever [plaintiffs] … inquired into Gorton's beliefs regarding his exposure to copyright liability at the time … [defendants] invoked privilege

…." *Id.* The court held that the defendants had waived privilege by putting good-faith at issue. *Id.* In the court's view, the defendants' unfair use of privilege to block inquiry into the critical facts justified waiver even though the defendants did not intend to present "explici[t]" evidence that they "relied upon … the advice of counsel." *Id.* at *2–3 (quoting *Leviton Mfg. Co. Inc.*, 2010 WL 4983183, at *3).

This case is nothing like *Arista Records.* For one, OpenAI has not made an affirmative good-faith defense. *See supra*, pp. __. For another, OpenAI has not sought to "rely"—either directly or indirectly—on its good faith belief about the lawfulness of its conduct *at the time it deleted the datasets.* Inquiry into OpenAI's reasons for deleting the datasets will not illuminate "the basis of [OpenAI's] understanding that [its] actions were legal" when it *used* the datasets to train GPT-3 and GPT-3.5 years earlier. *Arista Records*, 2011 WL 1642434, at *3 (quoting *Erie II*, 546 F.3d at 228). After all, willfulness or knowledge is measured at the time of the alleged infringement. *See Halo Elecs., Inc. v. Pulse Elecs., Inc.*, 579 U.S. 93, 105 (2016). And the plaintiffs have all the evidence they need on OpenAI's *use* of the datasets, including access to recovered copies of the datasets themselves and extensive discovery into how OpenAI used them in training. *See supra* at 1 & n.1.

V. CONCLUSION

For the foregoing reasons, OpenAI respectfully requests that the Court set aside Magistrate Judge Wang's November 24, 2025 Opinion and Order and deny class plaintiffs' motion to compel.

Respectfully submitted,

Dated: December 8, 2025

By: */s/ LISA S. BLATT*

LISA S. BLATT
WILLIAMS & CONNOLLY LLP
680 Maine Avenue SW
Washington, DC 20024
(202) 434-5000
lblatt@wc.com

Dated: December 8, 2025

By: */s/ R. JAMES SLAUGHTER*

R. JAMES SLAUGHTER
KEKER, VAN NEST & PETERS LLP
633 Battery Street
San Francisco, CA 94111
(415) 391-5400
rslaughter@keker.com

Dated: December 8, 2025

By: */s/ MARGARET GRAHAM*

MARGARET GRAHAM
LATHAM & WATKINS LLP
1271 Avenue of the Americas
New York, NY 10020
(212) 906-1200
margaret.graham@lw.com

Dated: December 8, 2025

By: */s/ ROSE LEE*

ROSE LEE
MORRISON & FOERSTER LLP
707 Wilshire Boulevard
Los Angeles, CA 90017
(213) 892-5355
roselee@mofo.com

Attorneys for Defendants
OPENAI, INC., OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, OPENAI HOLDINGS, LLC

* All parties whose electronic signatures are included herein have consented to the filing of this document, as contemplated by Rule 8.5(b) of the Court's ECF Rules and Instructions.