

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE:

OPENAI, INC.,
COPYRIGHT INFRINGEMENT LITIGATION

25-md-3143 (SHS) (OTW)

OPINION & ORDER

This Document Relates To:

All Putative Class Actions

23-cv-8292

23-cv-10211

25-cv-3291

25-cv-3482

25-cv-3483

SIDNEY H. STEIN, U.S. District Judge.

OpenAI¹ has objected to Magistrate Judge Ona T. Wang’s November 24, 2025 discovery order determining that OpenAI waived the attorney-client privilege with respect to communications in 2022 regarding the reasons OpenAI deleted the Books1 and Books2 datasets and communications in 2022 referencing Library Genesis. (See Dkt. No. 846 (“November 24 Order”); Dkt. No. 925.) The Court concludes pursuant to Federal Rule of Civil Procedure 72(a) that the November 24 Order is “clearly erroneous” or “contrary to law.” Simply put, OpenAI did not disclose material protected by the attorney-client privilege. In addition, OpenAI did not change its position regarding whether legal advice concerning the deletion of the Books1 and Books2 datasets is privileged, so waiver is not an appropriate sanction on the theory that OpenAI made a “moving target” of its privilege assertions. And OpenAI did not put its state of mind at issue merely by denying the Consolidated Class Action Complaint’s allegations that

¹ This Order refers to defendants OpenAI, Inc., OpenAI OpCo LLC, OpenAI GP, LLC, OpenAI, LLC, OpenAI Global LLC, OAI Corporation, OpenAI Holdings, LLC, OpenAI Startup Fund I LP, OpenAI Startup Fund GP I LLC, and OpenAI Startup Fund Management LLC collectively as “OpenAI.”

OpenAI willfully infringed Class Plaintiffs'² copyrights, so OpenAI did not effect an "at-issue" privilege waiver. The Court therefore sets aside the November 24 Order.

I. BACKGROUND

The Court assumes familiarity with the facts underlying this dispute and will repeat only those facts necessary to explain this Order.

In 2018, an OpenAI employee downloaded pirated copies of books from Library Genesis ("LibGen"), a website Class Plaintiffs have described as a "notorious shadow library" of digital copies of books. (See November 24 Order at 1.) Using the downloaded copies of books from LibGen, OpenAI created two datasets it employed to train the GPT-3 and GPT-3.5 large language models. It denominated those datasets "Books1" and "Books2." In mid-2022, OpenAI deleted the Books1 and Books2 datasets.

In March 2024, OpenAI's outside counsel informed Class Plaintiffs in a letter that concerned its document collection efforts for discovery in this litigation that "[w]e . . . understand that the use of [B]ooks1 and [B]ooks2 for model training was discontinued in late 2021, after the training of GPT-3 and GPT-3.5, and those datasets were then deleted in or around mid-2022 *due to their non-use.*" (See Case No. 23-cv-8292, Dkt. No. 106-4 ("March 2024 OpenAI Letter") at 3 (emphasis added).)

The following month, in opposition to a motion by Class Plaintiffs, OpenAI repeated that "the Books1 and Books2 datasets were deleted *due to non-use* before any litigation had been filed against OpenAI" and noted that OpenAI was attempting to locate any existing copies of the datasets. (Case No. 23-cv-8292, Dkt. No. 145 ("April 2024 OpenAI Letter Brief") at 3 (emphasis added).)³

Class Plaintiffs conducted a Rule 30(b)(6) deposition of Michael Trinh, Associate General Counsel for litigation and discovery at OpenAI, on January 29, 2025. (November 24 Order at 4; *see also* Dkt. No. 413-8 ("Jan. 2025 Trinh Dep.")). At that deposition, Trinh testified that "[t]he decision [to delete the Books1 and Books2 datasets] involved legal counsel" at OpenAI. (Jan. 2025 Trinh Dep. at 57–58.) Counsel for OpenAI then instructed Trinh not to answer any further questions about the

² This Order refers to plaintiffs in Case Nos. 23-cv-8292, 23-cv-10211, 25-cv-3291, 25-cv-3482, and 25-cv-3483 collectively as "Class Plaintiffs."

³ OpenAI was ultimately able to recover versions of the Books1 and Books2 datasets and it has produced them to Class Plaintiffs (*see* Dkt. No. 926 at 1 n.1), although Class Plaintiffs contend that the recovered versions of the datasets are incomplete (*see* Dkt. No. 996 ("Opposition") at 5).

deletion of Books1 and Books2 on the basis of the attorney-client privilege. (*Id.* at 61:19–62:3.)

In April 2025, based on OpenAI’s assertion of privilege at the January 2025 Trinh deposition, Class Plaintiffs moved for an order declaring that “non-use” was not a privileged reason for the deletion of Books1 and Books2 or that any privilege over the deletion of Books1 and Books2 due to their “non-use” had been waived based on OpenAI’s prior disclosure of non-use as a reason for the deletion of the datasets. (November 24 Order at 4.) OpenAI opposed that motion on the basis that “[t]he reasons for the deletion are privileged” (Dkt. No. 53-3 at 9), although it also stated that it would not “block[] plaintiffs from exploring the question of nonuse of the data set to be [a] cause of the deletions” (Dkt. No. 109 (“May 27, 2025 Hearing Tr.”) at 70:4–6). Based on this representation, Judge Wang ordered a further Rule 30(b)(6) deposition of OpenAI’s corporate designee on this issue. (May 27, 2025 Hearing Tr. at 70–75.)

Class Plaintiffs then served a new Rule 30(b)(6) deposition notice on OpenAI on June 10, 2025, that included the following topic: “All reasons why each version of [B]ooks1 and/or [B]ooks2 was deleted, including ‘non-use.’” (November 24 Order at 5–6.) Three days later, OpenAI filed a notice purporting to “withdraw” the March 2024 OpenAI Letter and the April 2024 OpenAI Letter Brief and replace them with versions of those documents that struck mentions of “non-use” as a reason for the deletion of the Books1 and Books2 datasets. (*Id.* at 6.) On June 29, 2025, OpenAI objected to Class Plaintiffs’ new Rule 30(b)(6) deposition notice on the grounds that it called for privileged information and stated that “OpenAI will not advance any non-privileged reason for the deletion of the [Books1 and Books2 datasets] in connection with this litigation.” (*Id.*; Dkt. No. 413-10 at 8.)

A second Rule 30(b)(6) deposition of Michael Trinh was held on July 25, 2025. (*See* Dkt. No. 413-1 (“July 2025 Trinh Dep.”).) At that deposition, Trinh testified that OpenAI employees as well as outside counsel were “involved in the decision to delete” the Books1 and Books2 datasets. (*Id.* at 82–92.) Trinh also identified people involved in the decision and identified the period in which the decision was made. (*Id.* at 75, 82–92.) Trinh repeated that OpenAI would not advance any non-privileged reasons for the deletion of the Books1 and Books2 datasets in this litigation but, at the instruction of OpenAI’s outside counsel, refused to answer questions about “the nonprivileged facts regarding . . . the reasons OpenAI deleted the Books1 and Books2 datasets.” (*Id.* at 102, 107.)

Five days after the July 2025 Trinh deposition, counsel for OpenAI informed Class Plaintiffs by email that “OpenAI does not believe that there are non-privileged reasons for the deletion [of Books1 and Books2].” (November 24 Order at 7; Dkt. No. 413-11 at 1.) The next day, Class Plaintiffs renewed their motion to compel discovery

regarding the deletion of the Books1 and Books2 datasets, contending that OpenAI had waived the attorney-client privilege by disclosing “non-use” as a reason for the deletion or, alternatively, that the crime-fraud exception to the attorney-client privilege applied. (November 24 Order at 8.) In opposition, OpenAI asserted that it had “consistently explained” that “there are *no* non-privileged reasons for the removal” of the Books1 and Books2 datasets. (Dkt. No. 428 at 1.) At oral argument on the motion before Judge Wang, OpenAI contended that the purpose of the March 2024 OpenAI Letter and April 2024 OpenAI Letter Brief was “to disclose that the data sets had been removed” and that the letters were not “intended to describe all the reasons for the removal.” (Dkt. No. 525 at 176.)

On November 24, 2025, Judge Wang found that OpenAI had waived the attorney-client privilege over communications in 2022 regarding the reasons for the deletion of the Books1 and Books2 datasets and communications in 2022 referencing LibGen. (November 24 Order at 27.)

The November 24 Order articulated three rationales for this finding. First, the November 24 Order assumed, based on OpenAI’s position that all reasons for the deletion of Books1 and Books2 were privileged, that “non-use” was a privileged reason for the deletion of the Books1 and Books2 datasets and found that OpenAI had impliedly waived the attorney-client privilege over communications related to the deletion of the datasets by disclosing “non-use” as a reason for the deletion. (November 24 Order at 15–17.) Second, the November 24 Order found that OpenAI had waived the attorney-client privilege by “making a moving target of its privilege assertions” by (1) “initially stat[ing], without equivocation, that Books1 and Books2 were deleted ‘due to their non-use’” and leaving “this assertion . . . on the docket for 15 months,” (2) “then assert[ing] that all the ‘reasons’ for the deletion of Books1 and Books2 were privileged,” (3) then claiming “that not all aspects of the ‘reasons’ for deletion were privileged, including the question of non-use as a ‘reason’ for deletion,” and (4) then retracting the March 2024 OpenAI Letter and April 2024 OpenAI Letter Brief and asserting again “that all the reasons for the deletion of Books1 and Books2 are privileged.” (*Id.* at 16.) Third, the November 24 Order found that OpenAI had waived the attorney-client privilege by putting its “good faith” at issue when it denied the allegations in the Consolidated Class Action Complaint that OpenAI had “willfully” infringed Class Plaintiffs’ copyrights. (*Id.* at 17–23.)

Separately, the November 24 Order determined that the crime-fraud exception to attorney-client privilege did not apply to the communications at issue because any communications about the deletion of the Books1 and Books2 datasets necessarily “must have taken place after the alleged crime of torrenting pirated books (*i.e.*, the acquisition of LibGen),” so the communications were not “made during and in

furtherance of a suspected crime.” (*Id.* at 26.) The November 24 Order also rejected Class Plaintiffs’ argument that the crime fraud exception applied on the basis of alleged spoliation simply because OpenAI deleted Books1 and Books2 “while facing ‘substantial legal uncertainty’ regarding the legality of its actions.” (*Id.* at 27.)

II. APPLICABLE LAW

A district judge may set aside a magistrate judge’s order on a non-dispositive matter only if the order “is clearly erroneous or is contrary to law.” Fed. R. Civ. P. 72(a); 28 U.S.C. § 636(b)(1)(A). “An order is ‘clearly erroneous’ only when the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed.” *Dass v. City Univ. of N.Y.*, No. 18-cv-11325, 2022 WL 16737028, at *2 (S.D.N.Y. Nov. 7, 2022) (quoting *JPMorgan Chase Bank, N.A. v. Reifler*, No. 11-cv-4016, 2016 WL 10570981, at *2 (S.D.N.Y. July 14, 2016)). “An order is contrary to law when it fails to apply or misapplies relevant statutes, case law, or rules of procedure.” *Id.* (quoting *Lan v. Time Warner, Inc.*, No. 11-cv-2870, 2016 WL 928731, at *1 (S.D.N.Y. Feb. 9, 2016)). “It is well-settled that a magistrate judge’s resolution of a non-dispositive matter should be afforded substantial deference” *JPMorgan*, 2016 WL 10570981, at *2 (quoting *Lan*, 2016 WL 928731, at *1).

III. ANALYSIS

Each of the bases for the November 24 Order’s finding that OpenAI waived the attorney-client privilege over communications concerning the deletion of the Books1 and Books2 datasets or referencing LibGen is clearly erroneous or contrary to law.

First, the statements in OpenAI’s letters that the Books1 and Books2 datasets were deleted “due to non-use” did not reveal any privileged information. Accordingly, OpenAI did not waive the attorney-client privilege by disclosing privileged information.

“The attorney-client privilege protects communications (1) between a client and his or her attorney (2) that are intended to be, and in fact were, kept confidential (3) for the purpose of obtaining or providing legal advice.” *United States v. Mejia*, 655 F.3d 126, 132 (2d Cir. 2011). “[A] party cannot partially disclose privileged communications or affirmatively rely on privileged communications to support its claim or defense and then shield the underlying communications from scrutiny by the opposing party.” *In re Grand Jury Proceedings*, 219 F.3d 175, 182 (2d Cir. 2000); *see also United States v. Bilzerian*, 926 F.2d 1285, 1292 (2d Cir. 1991).

OpenAI’s statements that Books1 and Books2 were deleted “due to non-use” do not reveal any legal advice from an attorney to OpenAI; the statements therefore did not disclose any privileged communications. Because the statements about deletion

“due to non-use” are not privileged, those statements cannot be a basis for finding that OpenAI waived the attorney-client privilege. Cf. *In re FirstEnergy Corp.*, 154 F.4th 431, 443 (6th Cir. 2025) (“[T]he disclosure of non-privileged material never waives the attorney-client privilege.”); *Aioissa v. Bank of Am., N.A.*, Case No. 10-cv-1275, 2011 WL 4026902, at *5 (E.D.N.Y. Sep. 12, 2011) (noting that selective disclosure waiver “is concerned with the selective disclosure of privileged material, not the disclosure of non-privileged material and redaction or withholding of privileged material”).

Second, OpenAI did not make a “moving target” of its privilege assertions such that a privilege waiver would be an appropriate sanction. OpenAI consistently asserted that any communications between OpenAI and its attorneys seeking or providing legal advice regarding the deletion of the Books1 and Books2 datasets were privileged communications. OpenAI has never asserted that its statements about deletion “due to non-use” — which were made in the context of letters regarding OpenAI’s document collection efforts — are protected by the attorney-client privilege, including when it attempted to remove those statements from the March 2024 OpenAI Letter and April 2024 OpenAI Letter Brief in June 2025. OpenAI’s inartful later formulation that “there are no non-privileged reasons” for the deletion of the Books1 and Books2 datasets (Dkt. No. 428 (emphasis omitted)) and improper instruction to Trinh not to answer a question seeking non-privileged facts related to the reasons for the deletion in the context of Trinh’s contentious July 2025 deposition are insufficient to demonstrate a “moving target” in OpenAI’s privilege assertions akin to cases in which courts in this district have found a privilege waiver on a “moving target” theory. See *In re Honeywell Int’l Income Sec. Litig.*, 230 F.R.D. 293, 299–300 (S.D.N.Y. 2003) (finding waiver as to particular documents, but not subject-matter waiver, where party originally asserted only attorney-client privilege and later asserted attorney work product protection over the documents as an attempt to “re-engineer [their] privilege log[] to align their privilege assertions with their legal arguments”); *SEC v. Yorkville Advisors, LLC*, 300 F.R.D. 152, 165 (S.D.N.Y. 2014) (finding privilege waiver where original privilege log contained insufficient detail and revised log asserted “claims of privilege that were never previously asserted” as to the documents); *Chealsea Hotel Owner LLC v. City of New York*, 21-cv-3982, 2024 WL 323216, at *2 (S.D.N.Y. Jan. 29, 2024) (declining to impose waiver despite party’s “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”), *aff’d*, 2025 WL 880530 (S.D.N.Y. Mar. 21, 2025).

Third, OpenAI has not put its state of mind regarding the legality of its conduct at issue merely by denying Class Plaintiffs’ allegations of willful copyright infringement. In its Answer to the Consolidated Class Action Complaint’s allegations that OpenAI willfully infringed Class Plaintiffs’ copyrights, OpenAI simply denied the

allegations of willfulness. (*See, e.g.*, Dkt. No. 754 ¶¶ 121, 175, 189, 200, 211, 221, 226, 233, 242, 247, 253, 266, 276, 286, 301, 312.)

To effect an at-issue waiver, “a party must *rely* on privileged advice from his counsel to make his claim or defense,” such as by affirmatively asserting “a good faith or state of mind defense.” *In re County of Erie*, 546 F.3d 222, 229 (2d Cir. 2008). In this case, OpenAI has not asserted a good faith defense. The November 24 Order nevertheless found that OpenAI has put its state of mind at issue because “[f]or OpenAI to deny that it willfully infringed class plaintiffs’ copyrighted works is to argue that it acted in good faith.” (November 24 Order at 18–19.) That conclusion is contrary to law.

There is a distinction between a copyright defendant merely denying allegations of willfulness—on which a plaintiff bears the burden of proof—and a copyright defendant affirmatively asserting its good faith belief that its actions were lawful. This distinction can have important consequences—for example, a finding of “willful” infringement permits a court to award increased statutory damages of up to \$150,000 per work, while a successful good faith defense permits the court to limit damages to \$200 per infringement. 17 U.S.C. § 504(c)(2). The U.S. Court of Appeals for the Second Circuit has highlighted that a finding of a lack of willfulness is not equivalent to a finding of innocent intent or good faith. *Fitzgerald Publ’g Co. v. Baylor Publ’g Co.*, 807 F.2d 1110, 1115 (2d Cir. 1986) (“It is plain that ‘willfully’ infringing and ‘innocent intent’ are not the converse of one another. Thus, it is possible in the same action for a plaintiff not to be able to prove a defendant’s willfulness, and, at the same time, for the defendant to be unable to show that it acted innocently.”); *see also Nat’l Football League v. PrimeTime 24 Joint Venture*, 131 F. Supp. 2d 458, 476 (S.D.N.Y. 2001) (“‘Innocent’ intent, however, is more than just the absence of willfulness.”); 5 Melville B. Nimmer & David Nimmer, *Nimmer on Copyright* § 14.04[B][3][a] (2026).

OpenAI has not asserted a good faith defense merely by denying Class Plaintiffs’ allegations of willfulness, and OpenAI has represented that it will not offer evidence of its state of mind regarding the legality of its conduct to rebut Class Plaintiffs’ allegations of willfulness at trial. (Jan. 16, 2026 Hearing Tr. at 32:11–17 (“We do not intend to defend [the action] on a good-faith belief that what we were doing is legal. We intend to defend it on other grounds, like fair use, no liability, and so on. . . . [W]e are not presenting a good-faith defense.”), 43:1–2 (“[W]e did elect, before Judge Wang, not to present a state-of-mind defense.”).)⁴ This case is therefore unlike the cases relied upon

⁴ The November 24 Order offered OpenAI the opportunity to avoid an at-issue privilege waiver if it “explicitly and unambiguously states its intention to forego asserting its good faith defense and any factual arguments in opposition to Class Plaintiffs’ allegations of willfulness.” (November 24 Order at

by Class Plaintiffs in which the defendants sought to use affirmative evidence of their good faith belief in the legality of their actions to defeat elements of the plaintiffs' cases.⁵ OpenAI thus has not effected an at-issue waiver of the attorney-client privilege by affirmatively relying on materials over which it asserts privilege to make its defense. *In re County of Erie*, 546 F.3d at 229.

Finally, Judge Wang's conclusion—after her *in camera* review of a sample of relevant documents—that the crime-fraud exception does not apply to the communications concerning the deletion of Books1 and Books2 was well founded. Judge Wang reasonably concluded that Class Plaintiffs failed to demonstrate probable cause that privileged communications regarding the deletion of Books1 and Books2 were made in furtherance of and were intended to conceal criminal copyright infringement. *See In re Grand Jury Subpoenas Dated Sep. 13, 2023*, 128 F.4th 127, 141–42 (2d Cir. 2025) (“A party seeking to invoke the crime-fraud exception must demonstrate that there is probable cause (1) that the client communication or attorney work product in question was *itself* in furtherance of the crime or fraud and (2) to believe that the particular communication with counsel or attorney work product was *intended* in some way to facilitate or to conceal the criminal activity.” (internal quotation marks and citation omitted)). This conclusion is especially well founded in light of the fact—as

23.) As set forth above, OpenAI has explicitly and unambiguously stated its intention to forego asserting a good faith or state of mind defense, as it must to avoid a waiver under *In re County of Erie* and *Bilzerian*. OpenAI need not, however, forego “any factual arguments in opposition to Class Plaintiffs’ allegations of willfulness” to avoid effecting an at-issue privilege waiver. As the Second Circuit explained in *In re County of Erie*, such a waiver occurs only when a defendant “rel[ies] on privileged advice from his counsel to make his claim or defense,” such as by “claim[ing] a good faith or state of mind defense.” 546 F.3d at 229 (emphasis omitted). OpenAI therefore may present factual arguments in opposition to Class Plaintiffs’ allegations of willfulness without waiving attorney-client privilege so long as OpenAI does not put its state of mind regarding the legality of its actions at issue by, for example, offering testimony that OpenAI employees believed their activities to be lawful.

⁵ *See Bilzerian*, 926 F.2d 1292–93 (finding an at-issue waiver based on the defendant’s proposed “testimony that he thought his actions were legal,” which “would have put his knowledge of the law and the basis for his understanding of what the law required in issue,” and noting a distinction between a defendant merely “denying criminal intent” and a defendant offering “testimony as to his good faith”); *Arista Records LLC v. Lime Group LLC*, No. 06-cv-5936, 2011 WL 1642434, at *1–2 (S.D.N.Y. Apr. 20, 2011) (finding at-issue waiver when defendants affirmatively asserted a “good faith belief in the lawfulness of their conduct”); *Okla. Firefighters Pension & Ret. Sys. v. Musk*, No. 22-cv-3026, 2025 WL 2982003, at *5–6 (S.D.N.Y. Oct. 23, 2025) (noting that “the assertion of a defense that a party had a good faith belief in the lawfulness of its conduct may impliedly waive the attorney-client privilege” and finding an at-issue waiver when defendants represented that they “intend[ed] to assert their good faith belief in the lawfulness of their actions (or omissions) after they first consulted with counsel”).

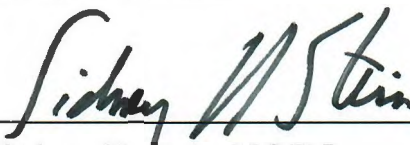
OpenAI notes—that OpenAI ultimately turned over to Class Plaintiffs copies of those datasets that it had recovered. (Dkt. No. 1004-1 at 11; Opposition at 5.) Judge Wang also reasonably concluded that Class Plaintiffs had failed to allege a sufficient factual basis for their contention that the crime fraud exception should apply on the theory that OpenAI’s destruction of the Books1 and Books2 datasets despite “substantial legal uncertainty” constituted spoliation.

IV. CONCLUSION

Accordingly, the Court sets aside the November 24 Order.

Dated: New York, New York
February 6, 2026

SO ORDERED:



Sidney H. Stein, U.S.D.J.