UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

---x

IN RE: :
                                                        :
OPENAI, INC.,   :   25-md-3143 (SHS) (OTW)
COPYRIGHT INFRINGEMENT LITIGATION,   :
                                                        :
                                                        :   **OPINION & ORDER RE: OPENAI'S**
This Document Relates To: :   **DELETION OF BOOKS1 AND BOOKS2**
23-CV-8292   :   **DATASETS AND PRIVILEGE RULINGS**
23-CV-10211   :   **(ECF NOS. 413, 428, 479, 481, 504, 505,**
24-CV-84   :   **615, 616)**
25-CV-3291   :
25-CV-3482   :
25-CV-3483   :
                                                        :
---x

ONA T. WANG, **United States Magistrate Judge**:

Class Plaintiffs seek a ruling on whether OpenAI must produce discovery and

communications related to the reasons for its deletion of the Books1 and Books2 datasets. It is

undisputed that in 2018, an OpenAI employee downloaded pirated copies of books from Library

Genesis ("LibGen"), (*see* ECF 413-1 at 151-168); (ECF 413-3 at 4),[1] a "notorious shadow library

which has been repeatedly ordered shut down by this court due to piracy." (ECF 368 at 1 n. 1).

See also Elsevier Inc. v. Sci-Hub, 15-CV-4282 (RWS), 2017 WL 3868800 (S.D.N.Y. 2017) (enjoining

Sci-Hub, Alexandra Elbakyan, and The Library Genesis Project from "unlawful access to, use,

reproduction, and/or distribution of Elsevier's copyrighted works"). From these downloaded

LibGen books, Class Plaintiffs assert that OpenAI created two datasets: "LibGen1" and

"LibGen2," which were later renamed to "Books1" and "Books2." (*See* ECF 413 at 2). In the

process of seeking discovery on the Books1 and Books2 datasets, Class Plaintiffs discovered,

---

[1] Unless otherwise indicated, all references to the docket refer to 25-MD-3143.

and it is undisputed that, in 2022, approximately one year before any of the actions in this

multidistrict litigation commenced, OpenAI deleted the Books1 and Books2 datasets. (ECF 413-

1 at 133). At the time, OpenAI asserted that the datasets were deleted due to "non-use." (23-

CV-8292, ECF 106-4). These are the only training datasets that, according to OpenAI, have ever

been deleted. (23-CV-8292, ECF 368-1 at 18). Then, when Class Plaintiffs sought discovery about

the reasons for the deletion of the Books1 and Books2 datasets, OpenAI asserted attorney-

client privilege. OpenAI's position on whether the reasons for the deletion are privileged has

shifted several times. Class Plaintiffs now assert that OpenAI's attorney-client privilege has

been waived by OpenAI's disclosure of privileged communications (e.g., selective voluntary

disclosure) and by putting the reasons for the deletion at issue through their affirmative

defense of innocent/non-willful infringement (e.g., at-issue waiver), or that the crime-fraud

exception applies to these communications. (*See* ECF 413).

For the following reasons, Class Plaintiffs' motion is **GRANTED in part, DENIED in part**.

I.   BACKGROUND

The Court assumes familiarity with the facts of this case and summarizes below the

factual and procedural history relevant to the pending motion.

A.   March—April 2024: OpenAI Asserts That the Books1 and Books2 Datasets
     Were Deleted In 2022 Due to "Non-Use"

On March 22, 2024, Mr. Gratz, outside counsel for OpenAI, sent Plaintiffs in the *Authors*

Guild (23-CV-8292) and *Alter* (23-CV-10211) actions an email with an update regarding the

Books1 and Books2 datasets. (*See* 23-CV-8292, ECF 106-4). In that email, Mr. Gratz informed

opposing counsel, apparently for the first time,[2] that the Books1 and Books2 datasets were discontinued as training datasets in late 2021, and that they had been deleted in mid-2022 "due to their non-use." (*Id.*).

Then, on April 12, 2024, Plaintiffs in the *Authors Guild* and *Alter* actions filed a motion to compel seeking, among other things, the identities of the two former OpenAI employees who were purportedly responsible for the deletion of the Books1 and Books2 datasets. (23-CV-8292, ECF 106). OpenAI filed a response on April 16, 2024, that unambiguously stated the reason for the deletion of the Books1 and Books2 datasets:

> First, Plaintiffs' motion omits key information that OpenAI had previously provided to them: namely, that the Books1 and Books2 datasets were deleted due to <u>non-use</u> before any litigation had been filed against OpenAI....

(23-CV-8292, ECF 145 at 3) (emphasis added). OpenAI also asserted in its opposition that it had agreed to supplement its interrogatory responses to provide Plaintiffs with the identities of the two former OpenAI employees involved in the deletion of Books1 and Books2. (*Id.*).

The *Authors Guild* and *Alter* actions were referred to me for general pretrial management on August 6, 2024. (ECF 177). The Court heard oral argument at the September 12, 2024, discovery status conference, where the parties agreed that the issue of the identities of the two former OpenAI employees had been resolved. (*See* 23-CV-8292, ECF 203 at 16). Accordingly, I denied the Plaintiffs' motion as moot. (*See* 23-CV-8292, ECF 202 at 2).

---

[2] Plaintiffs in the NY Authors Actions indicated that Mr. Gratz's March 22 email, which was sent while the parties were engaging in discovery regarding the training of OpenAI's LLMs, was the first time they learned that OpenAI had destroyed the Books1 and Books2 datasets. (23-CV-8292, ECF 106 at 3).

B. January 2025: OpenAI Asserts During the Deposition of Michael Trinh
   That the Reasons for the Deletion of the Books1 and Books2 Datasets Are
   Privileged

At the October 30, 2024, discovery status conference, I encouraged the parties to utilize

early 30(b)(6) depositions to help narrow the scope of document discovery. (23-CV-8292,

ECF 256 at 8). In accordance with that guidance, on January 13, 2025, the *Authors Guild* and

Alter Plaintiffs filed a motion seeking a five-hour 30(b)(6) deposition on the topic of the

location, storage, and retention of, among other things, training data. (23-CV-8292, ECF 308).

The Court granted the request on January 21, 2025, and directed the parties to conduct a

custodial 30(b)(6) deposition. (ECF 320).

Michael Trinh, Associate General Counsel for litigation and discovery at OpenAI, was

deposed on January 29, 2025, during which he was asked about the reasons for the deletion of

the Books1 and Books2 datasets. (*See generally* ECF 413-8). Nicholas Goldberg, counsel for

OpenAI, instructed Mr. Trinh not to answer because the question called for privileged

information, and Mr. Trinh stated that the decision to delete Books1 and Books2 "was made

with a consultation with the company attorneys, and I can't answer further without revealing

privileged information." (ECF 413-8 at 18-19).

C. May 20, 2025: OpenAI Asserts That All Reasons for the Deletion of
   Books1 and Books2 Are Privileged

Based on OpenAI's assertion of privilege at the January 29 deposition of Mr. Trinh, the

Authors Guild and *Alter* Plaintiffs filed a motion on April 1, 2025, seeking, *inter alia*, an order

deeming that "non-use" was not a privileged reason for the deletion of the Books1 and Books2

datasets, or, in the alternative, that any privilege over "non-use" had been waived based on

OpenAI's prior representations to the Court, or that the crime-fraud exception to attorney-

client privilege applies. (*See* 23-CV-8292, ECF 368). In advance of the May 27, 2025, status

conference, OpenAI asserted in the parties' joint chart that "[t]he reasons for the deletion [of

the Books1 and Books2 datasets] are privileged." (ECF 53-3 at 9).

> ### D. May 27, 2025: OpenAI Asserts In Court That Not All Aspects Regarding the Deletion of Books1 and Books2 Are Privileged

Six months ago, the Court heard oral argument on this issue. (May 27, 2025, Tr.,

ECF 109). At the May 27 conference, counsel for OpenAI stated: "[w]e are not asserting

privilege and <u>have not blocked plaintiffs from exploring the question of [non-use] of the</u>

<u>[dataset] to be [the] cause of the deletions</u>," (May 27 Tr. at 70) (emphasis added), and

explained further that "just because there may be an element of this issue that is not privileged,

[that] doesn't mean that every aspect of the issue is not privileged." (*Id.*). Based on counsel's

representation in court that "not all" aspects regarding the deletion of Books1 and Books2 were

privileged, I directed: (1) Plaintiffs to get discovery on the non-privileged aspects of the deletion

of Books1 and Books2; (2) the parties to meet and confer on what information regarding "non-

use" is not privileged; and (3) the parties to conduct an additional 30(b)(6) deposition regarding

the same. (May 27 Tr. at 70-71); (ECF 78 at 4) ("Class Plaintiffs' motion … regarding OpenAI's

assertion of privilege with respect to the deletion of Books1 and Books2 datasets is **DENIED as**

premature in light of OpenAI's representations on the record… Class Plaintiffs may also conduct

another 30(b)(6) deposition and are directed to serve a new 30(b)(6) notice on OpenAI

regarding the deletion of Books1 and Books2 and the reasons for those deletions."). Plaintiffs

served a new 30(b)(6) deposition notice on OpenAI on June 10, 2025. (ECF 413-10 at 1).

E.   June 13, 2025: OpenAI "Retracts" Prior Statements Regarding "Non-Use"

On June 13, 2025, <u>nearly 15 months</u> after OpenAI's initial representation that the

Books1 and Books2 datasets were deleted "due to non-use," but 3 days after Class Plaintiffs'

new 30(b)(6) notice, OpenAI filed a notice purporting to withdraw and replace the March 22,

2024, email from Mr. Gratz (23-CV-8292, ECF 143-4) and OpenAI's April 16, 2024, letter (23-CV-

8292, ECF 145). (*See* ECF 188). The proposed documents to replace the March 22, 2024, email

and April 16, 2024, letter struck all mentions of the Books1 and Books2 datasets being deleted

"due to non-use," but were otherwise identical to the originals. (*See* ECF Nos. 188-2, 188-4).

F.   June 29, 2025: OpenAI Asserts In Response to Class Plaintiffs' 30(b)(6)
Notice That It Will Not Advance Any "Non-Privileged Reasons for the
Deletion" of Books1 and Books2

OpenAI served its objections to Plaintiffs' new 30(b)(6) deposition notice on June 29,

2025. (ECF 413-10 at 8). In response to topic two, which reads: "All reasons why each version of

books1 and/or books2 was deleted, including 'non-use,'" OpenAI objected on the grounds that,

inter alia, it called for or could be construed as calling for, privileged information, and further

stated "OpenAI will not advance any non-privileged reason for the deletion of the BOOKS1 AND

BOOKS2 DATASETS in connection with this litigation." (ECF 413-10 at 8).

G.   July 25, 2025: OpenAI Asserts At the 30(b)(6) Deposition of Mr. Trinh That
All Reasons for the Deletion of Books1 and Books2 Datasets Are
Privileged

At the subsequent deposition of Mr. Trinh on July 25, 2025, Mr. Trinh testified that

OpenAI had engaged in both written and oral communications regarding the deletion of Books1

and Books2. (ECF 413-1 at 97-97). Then, when Class Plaintiffs asked Mr. Trinh to discuss the

"nonprivileged facts regarding … [t]he reasons OpenAI deleted the Books1 and Books2," Mr.

Goldberg again instructed Mr. Trinh not to answer Class Plaintiffs' question about non-

privileged facts because it was "designed to elicit a potential privilege waiver." (*Id.* at 107).[3]

Mr. Trinh did, however, answer a number of questions about non-privileged facts about Books1

and Books2, including how OpenAI downloaded the datasets from LibGen; that the datasets

were used to train GPT-3 and GPT-3.5; that Books1 and Books2 were not being used to train

any models at the time that they were deleted; the members of OpenAI's technical staff

involved in the deletion; that the discussions regarding deletion took place in June-July of 2022;

how technical staff members of OpenAI actually performed the deletion of Books1 and Book2;

and that copies of Books1 and Books2 were recovered after deletion. (ECF 413-1 at 56-131).

H. July 30, 2025: OpenAI Asserts That All The Reasons For the Deletion of Books1 and Books2 Are Privileged

In a July 30, 2025, email, counsel for OpenAI then represented to Class Plaintiffs that

"the reasons for the deletion are privileged," and that "**OpenAI does not believe that there are**

non-privileged reasons for the deletion." (ECF 413-11 at 2) (emphasis added). In other words,

every reason for the deletion of Books1 and Books2 is privileged.

OpenAI also stated in the July 30 email that "OpenAI does not intend to introduce

evidence of the advice of counsel either affirmatively or in rebuttal to the Class Plaintiffs'

burden of proof on any claim." (ECF 413-11 at 2). This assertion appears to suggest that OpenAI

will not claim that its alleged good faith is based directly on the specific advice of counsel.

OpenAI did not at this point, or any point since then, assert that it intends to forego its good

faith defense.

---

[3] It is unclear how the disclosure of admittedly non-privileged facts could lead to a privilege waiver. *C.f. In re Omeprazole Patent Litigation*, 227 F.R.D. 227, 231 (S.D.N.Y. 2005) ("[I]t is well established that voluntary disclosure of information to a party outside the privilege waives the attorney-client privilege with respect to such information.").

I.   August 2025: OpenAI Asserts That It Has "Consistently Maintained" That
     the Reasons for the Deletion of Books1 and Book2 Are Privileged

Class Plaintiffs renewed their motion to compel discovery regarding the deletion of the

Books1 and Books2 datasets on July 31, 2025, arguing that OpenAI had waived privilege, or, in

the alternative, that the crime-fraud exception applied to this information. (ECF 413). OpenAI

filed an opposition on August 4, 2025, arguing that OpenAI "has consistently explained" that

"there are no non-privileged reasons for the removal" of the Books1 and Books2 datasets.

(ECF 428).

The Court again heard oral argument on this issue at the August 12, 2025, discovery

status conference. (*See* Aug. 12, 2025, Transcript, ECF 525 at 160-177). At the August 12

conference, OpenAI maintained that it has been "consistent entirely that the reasons for the

removal of those data sets are privileged." (*Id.* at 164). OpenAI also asserted that previous uses

of the phrase "deleted due to non-use" were simply imprecise language that was "intended to

disclose that the data sets had been removed… They were not some letters that were intended

to describe all the reasons for the removal." (*Id.* at 176).

I then directed Class Plaintiffs to provide a timeline detailing OpenAI's positions on the

applicability of privilege to the reasons for the deletion of the Books1 and Books2 datasets,

(*id.* at 177), which Class Plaintiffs filed on August 22, 2025, (ECF Nos. 479, 481). OpenAI filed a

letter response on August 29, 2025. (ECF Nos. 504, 505). After reviewing the parties'

submissions, I directed the parties to file supplemental briefing on the issue of whether OpenAI

had waived privilege as to the reasons for the deletion of the Books1 and Books2 datasets.

(ECF 594). The parties filed their supplemental briefs on October 1, 2025. (*See* ECF Nos.

615, 616).

J.   The Court's October 1, 2025, Decision on OpenAI's Assertion of Privilege

On October 1, 2025, the Court issued an Opinion and Order (the "October 1 O&O")

regarding OpenAI's assertion of attorney-client privilege over 20 disputed documents that were

reviewed *in camera*. (ECF 618). In the October 1 O&O, I held that several documents containing

Slack messages between various OpenAI employees in 2022 regarding the deletion of LibGen

(e.g., the Books1 and Books2 datasets) were <u>not privileged</u> and ordered their production.[4] (*Id.*).

Also in the October 1 O&O, because of their relation to the issues in the instant motion, I

reserved my decision on the applicability of attorney-client privilege to (1) the portions of Log

Nos. 14 and 15 that appeared to be privileged and which concerned the deletion of LibGen data

and (2) the withheld documents under Log Nos. 17 and 18. (*Id.* at 10).

II.   **OPENAI'S REMAINING COMMUNICATIONS REGARDING THE DELETION OF THE BOOKS1 AND BOOKS2 DATASETS**

First, the Court will assess the applicability of attorney-client privilege to the remaining

disputed documents (Log Nos. 14, 15, 17, and 18) from Class Plaintiffs' prior motion to compel

that were submitted directly to the Court for *in camera* review. (*See* ECF Nos. 388, 412, 618).

Generally, parties may obtain discovery on any nonprivileged matter that is relevant to

any party's claim or defense. Fed. R. Civ. P. 26(b)(1). <u>Communications</u> that are protected by the

attorney-client privilege, however, are shielded from discovery. *Hayden v. Int'l Bus. Machs.*

Corp., 21-CV-2485 (VB) (JCM), 2023 WL 4622914, at *3 (S.D.N.Y. July 14, 2023).[5] "Attorney-

---

[4] The documents OpenAI was ordered to produce regarding the deletion of Library Genesis are Log Nos. 14, 15 and 16. (ECF 618 at 7-9).

[5] Attorney-client privilege does not protect facts from disclosure. *See Astra Aktiebolag v. Adrx Pharmas., Inc.*, 208 F.R.D. 92, 103 (S.D.N.Y. 2002). *See also Gucci Am., Inc. v. Guess?, Inc.*, 271 F.R.D. 58, 70 (S.D.N.Y. 2010) ("[P]rivilege only protects disclosure of communications; it does not protect disclosure of the underlying facts by those who communicated with 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 purposes of obtaining or providing legal advice.'" *Id.* (quoting *U.S. v. Mejia*, 655 F.3d 126, 132 (2d Cir. 2011)) (emphasis added). Where, as here, a case contemplates claims under federal law, federal law governs assertions of privilege. *Fresh Del Monte Produce, Inc. v. Del Monte Foods, Inc.*, 13-CV-8997 (JPO) (GWG), 2015 WL 3450045, at *2 (S.D.N.Y. May 28, 2015).

Attorney-client privilege only applies where the communication in question is "primarily or predominantly of legal character." *Egiazaryan v. Zalmayev*, 290 F.R.D. 421, 428 (S.D.N.Y. 2013). "Fundamentally, legal advice involves the interpretation and application of legal principles to guide future conduct or to assess past conduct." *In re County of Erie*, 473 F.3d 413, 419 (2d Cir. 2007). Where a party only asserts privilege over a limited portion of a document, courts in this Circuit "routinely consider the predominant purpose of [the] isolated portion." *In re Google Digital Advertising Antitrust Lit.*, 21-MD-3010 (PKC), 2023 WL 196146, at *2 n. 1 (S.D.N.Y. Jan. 17, 2023). "The burden is on a party claiming the protection of a privilege to establish those facts that are the essential elements of the privileged relationship, and they "must do so by competent and specific evidence, rather than by conclusory or *ipse dixit* assertions." *Hayden*, 2023 WL 4522914, at *3.

Log No. 14[6] contains Slack messages between Adam Nace, Alex Paino, Bob McGrew, Matt Knight, Nick Ryder, Lilian Weng, Jeff Wu, Jakub Pachocki, Jason Kwon, and Che Chang in a Slack channel titled "excise-libgen" or "project-clear," where these individuals discussed

---

[6] The beginning Bates Number associated with this document is OPCO_NDCAL_1615282.

deleting Library Genesis.[7] OpenAI withheld all of the Slack messages in this channel. In the October 1 O&O, I held that the vast majority of these communications were not privileged because they were "plainly devoid of any request for legal advice and counsel [did] not once weigh in." (ECF 618 at 7) (internal quotation marks omitted). Moreover, the only participation by Jason Kwon in those communications was to change the title of the chat from "excise-libgen" to "project-clear," which was also not a privileged communication because titles of documents are not privileged, notwithstanding the privileged nature of the contents of the document, unless the title itself, without reference to the document, contains privileged information. (*Id.*). The only communications from Log No. 14 that contain potentially privileged communications are (1) the message on page three (OPCO_NDCAL_1615284) from Bob McGrew to Jason Kwon and Matt Knight, which can be reasonably construed as a direct request for legal advice regarding Alex Paino's message starting with "general q," and (2) the response to Bob McGrew from Alex Paino. (*See* ECF 618 at 8).

As to Log No. 15,[8] which also contains Slack messages regarding the deletion of Library Genesis, the final message from Jason Kwon responding to Jeff Wu appears to be providing legal advice and is thus privileged. The remaining messages in this document are not privileged for the reasons discussed in the October 1 O&O. (*See* ECF 618 at 9).

Log No. 17.[9] contains Slack messages between Alex Paino, Matt Knight, Jeff Wu, Bob McGrew, Nick Ryder, Lilian Weng, and Jason Kwon regarding the deletion of Library Genesis.

---

[7] The names of any individual acting as an attorney for OpenAI at the time of the communications discussed below will be underlined.

[8] The beginning Bates Number associated with this document is OPCO_NDCAL_1615286.

[9] The beginning Bates Number associated with this document is OPCO_NDCAL_1615288.

Because these messages show a clear back-and-forth dialogue between various OpenAI

employees and Jason Kwon regarding legal issues and advice, the predominant purpose here is

to seek and provide legal advice, and these messages are privileged. *Hayden*, 2023 WL

4522914, at *3.

Log No. 18[10] is a list of links to internal OpenAI files related to Library Genesis. OpenAI

asserts that this document was attached to a message sent in the "excise-libgen"/"project-

clear" Slack channel and is thus privileged. (ECF 509-1 at 7). As I previously held in the October

1 O&O, the entirety of the Slack channel and all messages contained therein is not privileged

simply because it was created at the direction of an attorney and/or the fact that a lawyer was

copied on the communications. (ECF 618 at 6) (assessing privilege claims for individual

messages) (citing *In re Rivastigmine Patent Litigation*, 237 F.R.D. 69, 80 (S.D.N.Y. 2006); *Hayden*,

2023 WL 4522914, at *4)).[11] OpenAI's briefing does not make clear to which message the

document at Log No. 18 is attached. However, the final message of Log No. 17

(OPCO_NDCAL_1615291) appears to contain a file attachment, and the text of the message

indicates that the attached file is likely the list of links to internal OpenAI files found at Log No.

18. But the fact that the communications in Log No. 17 are privileged does not necessarily

confer privilege over the attachment itself.

"Attachments which do not, by their content, fall within the realm of the privilege

cannot become privileged by merely attaching them to a communication with an attorney."

Astra Aktiebolag v. Adrx Pharmas., Inc., 208 F.R.D. 92, 103 (S.D.N.Y. 2002) (internal quotation

---

[10] The beginning Bates Number associated with this document is OPCO_NDCAL_1615292.

[11] As also discussed in the October 1 O&O, other courts have taken the same approach to individual emails in an email chain. (*See* ECF 618 at 6 n. 10) (collecting cases).

marks omitted). *See also Roc Nation LLC v. HCC Int'l Ins. Co., PLC*, 2020 WL 1698928, at *2

(S.D.N.Y. Apr. 8, 2020) ("As to the issue of attachments to emails, HCC cannot assert that

attachments are automatically privileged—even if attached to a privileged email—and must

produce such attachments unless they themselves contain or refer to legal advice."); *In re*

Omeprazole Patent Lit., M-21-81 (BSJ), MDL 1291, 2005 WL 818821, at *8 (S.D.N.Y. Feb. 18,

2005). [12] Thus, the mere fact that the document at Log No. 18 was attached to a privileged

communication in Log No. 17 does not, by itself, confer privilege over Log No. 18. Looking to

the document itself, because a compilation of links does not contain or refer to legal advice,

explicitly or implicitly, this document (Log No. 18) is not privileged and must be produced. *See*

Roc Nation, 2020 WL 1698928, at *2.

III.   WAIVER

The Court next turns to whether OpenAI has waived attorney-client privilege over the

reasons for the deletion of the Books1 and Books2 datasets.

Courts in this district recognize several different types of privilege waivers. First, a party

waives attorney-client privilege if they disclose all or a significant part of a privileged

communication to a third party. *See In re Omnicom Group, Inc. Secs. Lit.*, 233 F.R.D. 400, 413

(S.D.N.Y. 2006). Such disclosure may also waive privilege for other communications on the

same topic (i.e., subject matter waiver). *In re Actos Antitrust Lit.*, 628 F. Supp. 3d 524, 532

(S.D.N.Y. 2022). Subject matter waiver may be appropriate in circumstances "in which fairness

---

[12] Other courts in this Circuit have come to similar conclusions. *See Eastman Kodak Co. v. Agfa-Gevaert N.V.*, 2006 WL 1495503, at *4 (W.D.N.Y. Apr. 21, 2006) ("Merely attaching a document to a privileged or protected document does not make the attached document privileged… [A] party objecting to the disclosure of a document with an attachment must prove that both the document and the attachment individually satisfy the requirements of the applicable privilege.") (internal quotation marks omitted).

requires a further disclosure of related, protected information, in order to prevent a selective

and misleading presentation of evidence to the disadvantage of the adversary." *Id.* (citing Fed.

R. Evid. 502(a), Advisory Committee Notes) (internal quotation marks omitted).

Second, privilege is waived where a party puts privileged communications "at issue."

See Scott v. Chipotle Mexican Grill, Inc., 67 F. Supp. 3d 607, 610 (S.D.N.Y. 2014). At-issue waiver

occurs where a party "asserts a claim that in fairness requires examination of protected

communications." *Id.* For example, in *U.S. v. Bilzerian*, the defendant raised a "good faith"

defense to securities fraud charges, but also claimed that conversations with counsel regarding

the same were privileged and shielded from disclosure. 926 F.2d 1285, 1292 (2d Cir. 1991). The

Second Circuit held that by putting "his knowledge of the law and the basis for his

understanding of what the law required [at] issue," the plaintiff impliedly waived attorney-

client privilege over communications "with counsel regarding the legality of his schemes" as

they were "directly relevant in determining the extent of his knowledge and, as a result, his

intent." *Id.* Courts continue to apply the "at-issue" waiver doctrine where "the assertion of a

good-faith defense involves an inquiry into state of mind" and the party asserting privilege

relies on privileged advice in making such a claim or defense.[13] *Chipotle Mexican Grill*, 67 F.

Supp. 3d at 610-11.

Other courts have found that attorney-client privilege may be waived where a party

changes its privilege claims over time in its privilege log. *See Chelsea Hotel Owner LLC v. City of*

New York, 21-CV-3982 (ALC) (RWL), 2025 WL 880530, at *4 (S.D.N.Y. Mar. 21, 2025). "'Neither

---

[13] Courts occasionally refer to the "at-issue" waiver as the "implied waiver." *See, e.g.*, *In re Omnicom Group*, 233 F.R.D. at 413 ("Implied waiver of the privilege will also be recognized—even in the absence of a party's disclosure of privileged material—when that party 'makes factual assertions, the truthfulness of which may be assessed only by an examination of the privileged communications or documents*.").

the Federal Rules of Civil Procedure nor the Local Civil Rules permit any party to make its

assertions of privilege a moving target.'" *Id.* (quoting *Sec. & Exch. Comm'n v. Yorkville Advisors,*

LLC, 300 F.R.D. 152, 166 (S.D.N.Y. 2014)). Further, "'[p]arties should not be permitted to re-

engineer privilege logs to align their privilege assertions with their legal arguments.'" *Id.*

(quoting *In re Honeywell Int'l Income Sec. Litig.*, 230 F.R.D. 293, 299-300 (S.D.N.Y. 2003)).

A. OpenAI Waived Privilege by Disclosing A "Privileged Reason" for the Deletion of the Books1 and Books2 Datasets

Through representations made to Class Plaintiffs and on the record, OpenAI has waived

privilege over "non-use" as a reason for the deletion of the Books1 and Books2 datasets.

As a preliminary matter, OpenAI's argument that the "reasons" for the deletion of the

Books1 and Books2 datasets are all privileged strains credulity. Only <u>communications</u> between

an attorney and their client are protected under the attorney-client privilege. *Hayden*, 2023 WL

4622914, at *3. As the Supreme Court has explained, attorney-client privilege does not protect

facts, even when they are shared with an attorney in otherwise privileged communications:

> The privilege only protects disclosure of communications; it does
> not protect disclosure of the underlying facts by those who
> communicated with the attorney: "[T]he protection of the privilege
> extends only to *communications* and not to facts. A fact is one thing
> and a communication concerning that fact is an entirely different
> thing. The client cannot be compelled to answer the question,
> 'What did you say or write to the attorney?' but <u>may not refuse to</u>
> <u>disclose any relevant fact within his knowledge merely because he</u>
> <u>incorporated a statement of such fact into his communication to</u>
> <u>his attorney</u>."

Upjohn Co. v. United States, 449 U.S. 383, 395-96 (1981) (internal citation omitted) (emphasis

added). It is unclear how "reasons"[14] can be privileged, even if such "reasons" were discussed

---

[14] OpenAI also makes the curious argument that the disclosure of "non-use" as a "reason" for the deletion on the public docket for 15 months does not waive privilege because the language "due to non-use" was not intended to describe a "reason" for the deletion of Books1 and Book2. This belated recharacterization cannot be squared with

with counsel in a privileged setting. Indeed, in the October 1 O&O, I held that many communications regarding the deletion of Books1 and Books2 were not privileged, and OpenAI did not object. (*See* ECF 618). Thus, if the "reasons" are not privileged, then there is no waiver of privilege because there is no privilege to waive, and OpenAI cannot refuse to testify about any of the reasons on the basis of attorney-client privilege.

Even if a "reason" like "non-use" <u>could</u> be privileged, OpenAI has waived privilege by making a moving target of its privilege assertions. OpenAI initially stated, without equivocation, that Books1 and Books2 were deleted "due to their non-use," (23-CV-8292, ECF 143-4), and this assertion was left on the docket for 15 months. They then asserted that <u>all</u> the "reasons" for the deletion of Books1 and Books2 were privileged, (ECF 413-11 at 2), and then claimed that <u>not all</u> aspects of the "reasons" for deletion were privileged, including the question of non-use as a "reason" for deletion. (May 27 Tr. at 70). This led directly to my Order for Class Plaintiffs to take more discovery. And it was weeks after my Order for discovery that OpenAI "retracted" its prior statements regarding non-use, (ECF 188), and asserted, again, that <u>all</u> the reasons for the deletion of Books1 and Books2 are privileged. (ECF 413-11 at 2). OpenAI's later assertions that all "reasons" are privileged compels a finding of waiver by disclosure. If all "reasons" are privileged, then "non-use," as a reason for the deletion, is privileged, but was disclosed multiple times.[15]

OpenAI's steadfast refusal to testify about "non-use" as a reason based on its potential to waive privilege. OpenAI's argument is also belied by the ordinary meaning of the words "due to." *Due To*, THE AM. HERITAGE COLL. DICTIONARY (5th ed. 2022) ("Because of."). Moreover, if "non-use" is not a "reason," then OpenAI cannot block inquiry into "non-use" on the basis of privilege. OpenAI's wielding of attorney-client privilege to block any inquiry into "non-use" clearly suggests that OpenAI considers it to be a privileged "reason," regardless of whether "reasons" as a category can even be privileged.

[15] OpenAI cites to no case law, from any district, to support the assertion underlying its argument that disclosure of privileged materials on a public docket can be cured months after the fact by filing a "retraction" of said disclosure.

Judge Lehrburger's reasoning in *Chelsea Hotel Owner LLC v. City of New York* is

instructive here. 21-CV-3982 (ALC) (RWL), 2024 WL 323216, at *2 (S.D.N.Y. Jan. 29, 2024),

report and recommendation adopted by 2025 WL 880530 (S.D.N.Y. Mar. 21, 2025). While Judge

Lehrburger ultimately held that the defendant's belated privilege assertions alone did not waive

privilege, the defendant there had consistently maintained that the documents were privileged

for one reason or another—the only shift was the type of privilege asserted. *Id.* Here, OpenAI

has gone back-and-forth on whether "non-use" as a "reason" for the deletion of Books1 and

Books2 is <u>privileged at all</u>. OpenAI cannot state a "reason" (which implies it is not privileged)

and then later assert that the "reason" is privileged to avoid discovery <u>after</u> Plaintiffs were

awarded discovery on that topic, regardless of whether such a change of position is "strategic

or a result of inadequate legal work." *Id.*[16] By virtue of disclosing a reason for the deletion of

Books1 and Books2, and then stating that all reasons are privileged, OpenAI has indeed

disclosed privileged material and waived privilege.

Accordingly, OpenAI has expressly waived privilege over attorney-client communications

regarding "non-use" as a reason for the deletion of Books1 and Book2.

B. OpenAI Also Waived Privilege by Putting Its "Good Faith" At Issue

OpenAI has also waived privilege over all <u>communications</u> regarding the "reasons" for

the deletion of Books1 and Books2 by putting its good faith and state of mind at issue.

In a copyright case, a court can increase the award of statutory damages up to $150,000

per infringed work if the infringement was willful, meaning the defendant "was actually aware

---

[16] In adopting Judge Lehrburger's Report & Recommendation, Judge Carter reiterated: "Neither the Federal Rules of Civil Procedure nor the Local Civil Rules permit any party to make its assertion of privilege a moving target." *Chelsea Hotel*, 2025 WL 880530, at *4. Nothing indicates that this general principle does not apply to a party's assertions of privilege in court or at deposition.

of the infringing activity" or the "defendant's actions were the result of reckless disregard for,

or willful blindness to, the copyright holder's rights." *See Broadcast Music, Inc. v. Prana*

Hospitality, Inc., 158 F. Supp. 3d 184, 197 (S.D.N.Y. 2016) (internal quotation marks omitted).

Where a defendant argues that any alleged infringing activity was "innocent," they must show

"(1) a subjective good faith belief in the innocence of its conduct that was (2) objectively

reasonable under the circumstances." *McDermott v. Kalita Mukul Creative Inc.*, 757 F. Supp. 3d

301, 311 (E.D.N.Y. 2024). As courts in this district have recognized, there is a fundamental

conflict where a party asserts a good faith defense based on advice of counsel but then blocks

inquiry into their state of mind by asserting attorney-client privilege:

> Plaintiffs are correct that a party may not assert that it believed its
> conduct was lawful, and simultaneously claim privilege to block
> inquiry into the basis for the party's state of mind or belief. Indeed,
> a party "cannot be permitted, on the one hand, to argue that it
> acted in good faith and without an improper motive and then, on
> the other hand to deny access to the advice given by counsel where
> that advice played a substantial and significant role in formulating
> actions taken by the defendant." Accordingly, "a party who intends
> to rely at trial on the advice of counsel must make a full disclosure
> during discovery; *failure to do so constitutes a waiver of the advice-
> of-counsel defense.*"

Arista Records LLC v. Lime Group LLC, 06-CV-5936 (KMW), 2011 WL 1642434, at *2 (S.D.N.Y.

Apr. 20, 2011) (internal citations omitted) (emphasis in original).

OpenAI's good faith and state of mind <u>are</u> at issue in this case.[17] In the Consolidated

Class Action Complaint ("CCAC"), Class Plaintiffs allege facts that support a reasonable inference

that OpenAI's infringement was willful. (*See, e.g.*, ECF 183 ¶¶ 76-312). In its recently filed

---

[17] Before the MDL was convened, the *Authors Guild* Plaintiffs alleged that OpenAI willfully infringed their
copyrighted materials, (*see* 23-CV-8292, ECF 69 ¶¶ 79-130),[17] and OpenAI repeatedly asserted that its conduct was
"innocent, [and] not willful." (*See, e.g.*, 23-CV-8292, ECF 75 at 49).

answer to the CCAC, OpenAI has artfully removed its good faith affirmative defense and key

words such as "innocent," "reasonably believed," and "good faith." (*See generally* ECF 754).[18]

However, upon closer inspection, OpenAI continues to dispute Class Plaintiffs' willfulness

allegations (i.e., argue that it acted innocently or in good faith). Class Plaintiffs allege that

OpenAI willfully infringed their copyrighted works, that whether OpenAI acted willfully is a

question of fact or law common to the purported class, and that OpenAI's conduct was and

continues to be willful. (*See, e.g.*, ECF 183 ¶¶ 121, 175, 189, 200, 211, 221, 226, 233, 242, 247,

253, 266, 276, 286, 301(d), 312). OpenAI unequivocally denies each and every one of these

factual allegations. (*See, e.g.*, ECF 754 ¶¶ 121, 175, 200, 211, 221, 226, 233, 242, 247, 253, 266,

276, 286, 301, 312). For OpenAI to deny that it willfully infringed Class Plaintiffs' copyrighted

works is to argue that it acted in good faith.

Because OpenAI continues to make factual assertions that its conduct was not willful,

and in the absence of any clear representations that OpenAI intends to waive this affirmative

defense or will not advance factual arguments opposing Class Plaintiffs' claims of willfulness,

OpenAI has put its alleged good faith and state of mind at issue.[19] And, because OpenAI has put

its state of mind at issue, it cannot simultaneously block discovery into evidence potentially

undermining the same by asserting attorney-client privilege. Contrary to OpenAI's argument

---

[18] OpenAI continues to assert such an affirmative defense in its Answer to the News Plaintiffs' complaint. OpenAI's twentieth and twenty-first affirmative defenses claim that OpenAI "reasonably believed that the acts alleged in the First Amended Complaint were not an infringement of copyright," and that "OpenAI did not know or have reasonable grounds to know that the acts alleged in the First Amended Complaint were an infringement of copyright." (*See* ECF 164 at 35-36) (emphasis added).

[19] In OpenAI's prior briefing on this issue, it suggests that "Plaintiffs may be trying to cast OpenAI's removal of the datasets as improper or nefarious. But these same plaintiffs' lawyers took the exact opposite position in *Bartz v. Anthropic*, arguing that it was wrongful for Anthropic to 'hoard' works it was no longer using." (ECF 504 at 2). But the reason why OpenAI deleted the datasets and its relevance to Plaintiffs' allegations of willfulness is a distinct issue from whether hoarding unlawfully obtained copyrighted works in an internal dataset is a copyright violation.

that Class Plaintiffs' willfulness theory is somehow limited to "OpenAI's creation and use of the

training datasets, [and] not their later removal by different employees years afterward,"

(ECF 504 at 2), it is not a stretch for Class Plaintiffs to posit that communications regarding the

reasons for deleting Books1 and Books2 could be probative of OpenAI's willfulness. Indeed, the

Court's *in camera* review of Log Nos. 14, 15, 17, and 18, discussed above, confirm that such

communications are likely probative of willfulness.

This outcome is not precluded by OpenAI's argument that it will not specifically rely on

any particular communications protected by the attorney-client privilege. Indeed, OpenAI's

wielding of attorney-client privilege to block <u>any inquiry</u> by Class Plaintiffs into OpenAI's state of

mind is improper. *Arista Records LLC v. Lime Group LLC* is instructive here. 2011 WL 1642434. In

Arista Records, plaintiff Arista Records ("Arista") brought secondary copyright infringement

claims against defendant LimeWire LLC ("LW"). *Id.* at 1. Following summary judgment in favor

of Arista, Arista moved to preclude LW from arguing at the trial on damages that LW had acted

in good faith (e.g., that their infringement was not willful). *Id.* A member of LW had previously

offered testimony on his state of mind—that he thought what he was doing was not a legal

risk—but LW invoked attorney-client privilege to "block <u>any inquiry</u> by [Arista] into [LW's]

alleged good faith belief that their conduct in operating [LW] was lawful." *Id.* (emphasis added).

As OpenAI has asserted here, LW argued that the Second Circuit's holding in *Bilzerian*

did not apply because LW would "not defend against assertions that they acted willfully or

engaged in fraudulent transactions by testifying that they relied upon … the advice of counsel,

or that their conduct was based on the advice of lawyers." *Id.* at 2. Like OpenAI, LW

misunderstood *Bilzerian* and subsequent case law. As Judge Wood explained:

> However, the *Erie* Court noted that the *Bilzerian* Court was correct
> in finding that, <u>if a defendant "asserted his good faith, the jury
> would be entitled to know the basis of his understanding that his
> actions were legal</u>." Further, a decision issued after *Erie* makes
> clear that: a party need not explicitly rely on advice of counsel to
> implicate the privileged communications. Instead, advice of
> counsel may be placed in issue where, for example, a party's state
> of mind, such as [their] good faith belief in the lawfulness of [their]
> conduct, is relied upon in support of a claim of defense…. <u>Because
> the legal advice that a party received may well demonstrate the
> falsity of its claim of good faith belief, waiver in these instances
> arises as a matter of fairness</u>.

Arista Records, 2011 WL 1642434, at *3 (quoting *Leviton Mfg. Co. Inc. v. Greenberg Traurig LLP*,

09-CV-8083 (GBD) (THK), 2010 WL 4983183, at *3 (S.D.N.Y. Dec. 6, 2010)) (emphasis added).

Courts in this district and others have routinely come to same conclusion. *See Oklahoma*

Firefighters Pension and Retirement System v. Musk, --- F. Supp. 3d ---, 2025 WL 2982003, at *4

(S.D.N.Y. Oct. 23, 2025) (collecting cases); *S.E.C. v. Honig*, 2021 WL 5630804, at *11 (S.D.N.Y.

Nov. 30, 2021) ("[C]ourts within this Circuit, relying on *Bilzerian*, have reaffirmed the broader

principle that forfeiture of the privilege may result when the proponent asserts a good faith

belief in the lawfulness of its actions, even without expressly invoking counsel's advice.")

(internal citations omitted).

OpenAI continues to assert that it did not willfully infringe Class Plaintiffs' copyrighted

works.[20] A jury is entitled to know the basis for OpenAI's purported good faith. It is immaterial

---

[20] OpenAI, bizarrely, cites to Judge Alsup's decision in *Bartz v. Anthropic PBC* to somehow suggest that
downloading pirated copies of books is lawful as long as they are subsequently used for training an LLM.
(*See* ECF 615 at 1 (citing *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007, 1026 (2025)). This is a gross
misunderstanding of Judge Alsup's decision, which states, in relevant part:

> This order doubts that any accused infringer could ever meet its burden of explaining why
> downloading source copies from pirate sites *that it could have purchased or otherwise
> accessed lawfully* was itself reasonably necessary to any subsequent fair use… Such piracy
> of otherwise available copies is inherently, irredeemably infringing <u>even if the pirated
> copies are immediately used for the transformative use and immediately discarded</u>….

whether OpenAI purports to rely on any specific communication(s) with counsel.[21] What

matters is that OpenAI has put its state of mind at issue, and OpenAI may not selectively use

attorney-client privilege to restrict Class Plaintiffs' <u>inquiry</u> into evidence concerning OpenAI's

purported good faith in this way. *See In re Actos*, 628 F. Supp. 3d at 532. And OpenAI has

repeatedly done so, even where Class Plaintiffs explicitly sought non-privileged facts and/or

reasons relating to the deletion of Books1 and Books2.[22] When Class Plaintiffs sought clarity on

---

Bartz, 787 F. Supp. 3d at 1025-26 (emphasis added). Contrary to its assertions, OpenAI's "new" reason for the deletion of Books1 and Books2 would seem to fall squarely into the category of activities proscribed by Judge Alsup.

[21] Judge Gorenstein dealt with a similar issue in *Oklahoma Firefighters Pension*. *See* 2025 WL 2982003. There, defendants in a securities class action asserted a good faith defense but stated that they were not "asserting an advice-of-counsel defense" or "rely[ing] on attorney-client privileged communications or documents in the[ir] defense." *Id.* at 2-4. Judge Gorenstein arrived at the same conclusion that the Court does here:

> Here, it would obviously be unfair for defendants to be permitted to offer evidence or argue that they had a good faith belief that their alleged conduct was lawful while at the same time barring plaintiff from examining the evidence that would most obviously put that claim to the test: the advice defendants' attorneys gave them about that very conduct. Accordingly, by stating their intention to offer such evidence, defendants waive attorney-client privilege as to their communications with counsel.

Id. at 6.

[22] At the second deposition of Mr. Trinh, counsel for OpenAI repeatedly blocked Class Plaintiffs' <u>inquiry</u> into the non-privileged facts regarding the reasons for the deletion of Books1 and Books2:

> By Mr. Smyser: Q. Will you tell me the nonprivileged facts regarding the reasons to delete the – the reasons OpenAI deleted the Books1 and Books2 datasets?
>
> Mr. Goldberg: Hang on. Again, because that is designed to elicit a potential privilege waiver, I will again ask you on the record, will Plaintiffs agree that if OpenAI provides nonprivileged facts in this regard it will not constitute a waiver of any privilege and execute a 502 stipulation.
>
> …
>
> Q. And I'm going to ask are there any nonprivileged facts that would contradict that statement?
>
> Mr. Goldberg: Counsel – Mr. Trinh, we've been over this now for the last 15 minutes. The Plaintiffs are not willing to say that if we provide facts or nonprivileged reasons for deletion of datasets that that information would not waive privilege. They are contending that the provision of such information

whether OpenAI would forego asserting any "nonprivileged facts" to contradict Class Plaintiffs'

possible future assertions that OpenAI's deletion of Books1 and Books2 may be evidence of

willfulness, Mr. Trinh refused to answer (at direction of counsel), and repeated his boilerplate

response that OpenAI would not advance any non-privileged reasons for the deletion of Books1

and Books2 in this case. (*See* ECF 413-1 at 113-116).

Thus, given the circumstances of this case, and as a matter of fairness to prevent the

selective and potentially misleading presentation of evidence, unless OpenAI 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,[23] OpenAI has waived

attorney-client privilege. *See Arista Records*, 2011 WL 1642434, at *3.

C. Scope of Waiver

Having found that OpenAI waived privilege through voluntary disclosure of privileged

communications and by putting legal advice at issue, the Court now turns to the scope of the

waiver. The appropriate scope of a privilege waiver is governed by the "fairness doctrine." *In re*

von Bulow, 828 F.2d 94, 101 (2d Cir. 1987). "The inquiry under this doctrine is to what extent is

disclosure necessary to prevent prejudice to a party and distortion of the judicial process that

---

> would waive privilege. And therefore, the provision of that information would
> cause OpenAI to lose its privilege over all of the communications relating to the
> reasons for deletion. For that reason, you can't answer that question.

(*See* ECF 413-1 at 107, 111-112).

[23] In *Oklahoma Firefighters Pension*, when prompted by Judge Gorenstein, the defendants stated that they
intended to assert a good faith defense "only if the Court determines that Defendants may do so without waiver of
the attorney-client privilege." 2025 WL 2982003, at *6. Upon finding that the defendants had waived privilege,
Judge Gorenstein accepted the "defendants' choice not to present that defense," and precluded them from
"offering argument or evidence that they had a good faith belief" in the lawfulness of their conduct." *Id.* OpenAI
has not yet made such a representation. However, if OpenAI were to make such an assertion, its choice would
likewise be accepted and enforced.

may be caused by the privilege-holder's selective disclosure during litigation of otherwise privileged information." *In re Keurig Green Mountain Single Serve Coffee Antitrust Litigation*, 14-MD-2542 (VSB) (HBP), 2019 WL 2724269, at *4 (S.D.N.Y. July 1, 2019) (internal quotation marks omitted). Where a party has waived privilege by putting privileged communications at issue "by virtue of his claims or defenses in the case, then a broader, so-called 'subject matter' waiver" is appropriate. *Bowne of New York City, Inc. v. AmBase Corp.*, 150 F.R.D. 465, 484 (S.D.N.Y. 1993) (quoting *In re von Bulow*, 828 F.2d at 102-03). Subject-matter waiver reaches "all privileged conversations regarding a particular subject." *In re von Bulow*, 828 F.2d at 103. "The scope of [a party's] waiver of the attorney-client and work-product privileges [is] a determination that lies within the sound discretion of the district court ...." *See Convolve, Inc. v. Compaq Comp. Corp.*, 00-CV-5141 (GBD), 2006 WL 2788234, at *2 (S.D.N.Y. Sept. 27, 2006).

Because OpenAI waived privilege by putting its good faith and state of mind at issue, I find that OpenAI has waived privilege over all communications in 2022 related to the reasons for the deletion of (1) the Books1 and Books2 datasets and (2) all internal references to LibGen. *See Bowne*, 150 F.R.D. at 484.

IV.   Crime-Fraud Exception

"The crime-fraud exception strips the privilege from attorney-client communications that relate to client communications in furtherance of contemplated or ongoing criminal or fraudulent conduct." *In re John Doe, Inc.*, 13 F.3d 633, 636 (2d Cir. 1994) (internal citations omitted). *See also Freedman v. Weatherford Intern. Ltd.*, 12-CV-2121 (LAK) (JCF), 2014 WL 3767034, at *5 (S.D.N.Y. July 25, 2014). "Thus, documents that 'would otherwise be protected from disclosure by the attorney-client privilege are not protected if they relate to communications in furtherance of a crime or fraud.'" *Zimmerman v. Poly Prep Country Day*

School, 2012 WL 2049493, at *6 (E.D.N.Y. June 6, 2012) (quoting *Duttle v. Bandler & Kass*, 127

F.R.D. 46, 53 (S.D.N.Y. 1989)). "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 … was *intended* in some way to facilitate or to conceal

the criminal activity." *In re Grand Jury Subpoenas Dated Sept. 13, 2023*, 128 F.4th 127, 141-42

(2d Cir. 2025).[24] Probable cause may be found where "a prudent person has a reasonable basis

to suspect the perpetration or attempted perpetration of a … fraud." *Chevron Corp. v. Donziger*,

11-CV-0691, 2013 WL 1087236, at *12 (S.D.N.Y. Mar. 15, 2013).

The crime-fraud exception is significantly broader than its titular terms of "crime" and

"fraud" may imply. *See Chevron Corp. v. Salazar*, 275 F.R.D. 437, 452 (S.D.N.Y. 2011) ("Case law

on the crime-fraud exception does not make perfectly clear what wrongdoing must be alleged,

and with what specificity, in order for the privilege to apply."). Clearly, the exception applies to

common-law fraud, and many courts have found the exception applies to "non-fraud

intentional torts." *See Specialty Minerals, Inc. v. Pleuss-Stauffer AG*, 98-CV-7775, 2004 WL

---

[24] Typically, a lower evidentiary threshold is required to conduct *in camera* review on allegations of crime-fraud than is needed to ultimately overcome attorney-client privilege. *See Amusement Industry, Inc. v.* Stern, 293 F.R.D. 420, 427 n. 4 (S.D.N.Y. 2013). When that lower threshold is met, *in camera* review is "an appropriate means for the court to determine whether allegedly privileged communications fall within the exception." *Zimmerman*, 2012 WL 2049493 at 9. Initially, I reviewed 20 documents *in camera* to determine the applicability of attorney-client privilege to those documents. That dispute also contained a passing request for *in camera* review based on the application of the crime-fraud exception, as does Class Plaintiffs' supplemental briefing. (*See* 23-CV-8292, ECF 368). The applicability of the crime-fraud exception was also raised at the August 12 conference. (*See* Aug. 12 Tr. at 163). Based on Class Plaintiffs' arguments at the August 12 conference, which arguments were bolstered by my *in camera* review of the 20 disputed documents, I determined that the lower threshold for *in camera* review for purposes of determining whether the crime-fraud exception applied was met, and further reviewed a subset of those documents for purposes of determining whether the crime-fraud exception applies or not. *See Amusement*, 293 F.R.D. at 427 n. 4 ("In order to review documents *in camera*, there need only be a 'showing of a factual basis adequate to support a good faith belief by a reasonable person that *in camera* review of the materials may reveal evidence to establish the claim that the crime-fraud exception applies.'").

42280, at *9 (S.D.N.Y. Jan. 7, 2004). Other courts have applied the exception to "misconduct

fundamentally inconsistent with the basis premises of the adversary system," *see Madanes v.*

Madanes, 199 F.R.D. 135, 148-49 (S.D.N.Y. 2001), and intentional torts that undermine the

adversarial system. *Chevron Corp.*, 275 F.R.D. at 453. Courts in other circuits have also held that

the crime-fraud exception applies to "a lawyer's unprofessional conduct," *Parrott v. Wilson*, 707

F.2d 1262, 1271 (11th Cir. 1983), and "bad faith litigation conduct," *Cleveland Hair Clinic, Inc. v.*

Puig, 968 F. Supp. 1227, 1241 (N.D. Ill. 1996)..[25]

Here, Class Plaintiffs present two theories in support of their claim that the crime-fraud

exception applies here. First, they assert that OpenAI's torrenting of pirated books is evidence

of criminal copyright infringement. (ECF 413 at 4). Class Plaintiffs alleged a similar theory in

Kadrey v. Meta Platforms, Inc., which Judge Hixson firmly denied. (N.D. Cal., 23-CV-3417,

ECF 627 at 2-3). The communications that Class Plaintiffs seek—those between OpenAI and

their counsel regarding <u>deleting</u> the Books1 and Books2 datasets—necessarily must have taken

place after the alleged crime of torrenting pirated books (i.e., the acquisition of LibGen). The

crime-fraud exception strips privilege from communications made during and in furtherance of

a suspected crime, not those that take place after. *See In re John Doe, Inc.*, 13 F.3d at 636. For

the same reasons outlined in Judge Hixson's opinion, Class Plaintiffs have failed to sustain their

burden under this theory. *See Kadrey* (N.D. Cal., 23-CV-3417, ECF 627 at 3) ("Nor do Plaintiffs

make any contention that, as a factual matter, the alleged crime was still being planned or was

ongoing at the time this document was created. Accordingly, Plaintiffs do not satisfy their

burden to show that the crime-fraud exception applies.").

---

[25] *See also Rambus, Inc. v. Infineon Technologies AG*, 220 F.R.D. 264, 283 (E.D. Va. 2004) (holding that
communications in furtherance of spoliation triggered the crime-fraud exception).

Second, Class Plaintiffs allege that the deletion of Books1 and Books2 "while facing 'substantial legal uncertainty' regarding the legality of its actions" constitutes misconduct at odds with the basic premises of the adversarial system. (ECF 413 at 5). At baseline, Class Plaintiffs have failed to sufficiently define the "misconduct" OpenAI has engaged in and how that conduct is "at odds" with the adversarial system, or why the caselaw in this Circuit justifies an extension of the crime-fraud exception to this alleged misconduct. *C.f. Rambus*, 220 F.R.D. at 281-83 (explaining how Fourth Circuit precedent justified extending the crime-fraud exception to spoliation). Moreover, Class Plaintiffs fail to allege a sufficient factual basis to find probable cause to believe that that the communications at issue were in furtherance of the types of misconduct that are recognized by Second Circuit precedent and/or that such communications were intended to facilitate or conceal such activities.

Accordingly, Class Plaintiffs have failed to show that the crime-fraud exception applies to these communications.

V.   CONCLUSION

For the foregoing reasons, Class Plaintiffs' motion to compel is **GRANTED in part** with respect to privilege waiver and **DENIED in part** with respect to the crime-fraud exception.

OpenAI is directed to produce forthwith the communications that the Court reviewed *in camera* (Log Nos. 14, 15, 17, and 18), and all other written communications with in-house counsel in 2022 regarding the reasons for the deletion of (a) the Books1 and Books2 datasets and (b) all internal references to LibGen that OpenAI has redacted or withheld on the basis of attorney-client privilege. OpenAI is further directed to log all written communications regarding the same to the extent they are not already on OpenAI's privilege log, and to identify on its

privilege log the specific communications that relate to the deletion of (a) and (b) above.

OpenAI's production must be completed by **December 8, 2025.**

Further, because OpenAI's 30(b)(6) deponent testified that OpenAI's counsel

participated in oral communications regarding the deletion of the Books1 and Books2 datasets

and OpenAI asserted attorney-client privilege over these communications, (*see* ECF 413-1 at 97-

98), Class Plaintiffs are entitled to depose OpenAI's in-house lawyers on such communications

and their personal knowledge after OpenAI has produced the written communications

described above. Given OpenAI's ongoing improper privilege assertions, Class Plaintiffs are

entitled to depositions of up to 2 hours for each OpenAI attorney who participated in such

communications in 2022, which will not count against the total deposition hours cap previously

set by the Court. OpenAI is directed to identify the OpenAI attorneys by **December 5, 2025,** and

make them available for these depositions no later than **December 19, 2025.**

The Clerk of Court is respectfully directed to close ECF 413.

SO ORDERED.

s/ Ona T. Wang

Dated: November 24, 2025   **Ona T. Wang**
      New York, New York   United States Magistrate Judge