

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

-----X

IN RE:	:	
	:	
OPENAI, INC.,	:	25-md-3143 (SHS) (OTW)
COPYRIGHT INFRINGEMENT LITIGATION,	:	
	:	
	:	<u>OPINION & ORDER</u>
This Document Relates To:	:	
All Actions	:	
	:	

-----X

ONA T. WANG, United States Magistrate Judge:

The Court is in receipt of ECF Nos. 727, 747 regarding the parties’ dispute as to Plaintiffs’ proposed Rule 30(b)(6) deposition topics. For the following reasons, Plaintiffs’ motion is

GRANTED.

I. BACKGROUND

On October 15, 2025, OpenAI filed a motion seeking a protective order as to Plaintiffs’ Rule 30(b)(6) notices, “which cumulatively contain 517 topics and subtopics.” (ECF 654). OpenAI’s motion indicated that after thorough meeting and conferring, OpenAI had agreed to provide testimony on 52 topics, but that disputes remained for an additional 59 topics that Plaintiffs identified as “priority disputes.” (*Id.* at 1-2). OpenAI seeks a protective order to limit its future 30(b)(6) testimony to the 52 topics for which OpenAI has already agreed to provide testimony. (*Id.* at 2). Plaintiffs filed a joint opposition on October 21, 2025, (ECF 688), arguing that OpenAI’s motion is improper because (1) OpenAI seeks to strike hundreds of topics that it did not even mention in its brief; (2) Microsoft agreed to ninety percent of the same topics for

which OpenAI now seeks a protective order; and (3) OpenAI's purported burden of preparing a witness is overblown because OpenAI is "inflat[ing] the number" of topics. (ECF 688).¹

The Court heard oral argument on this dispute at the October 29, 2025, discovery status conference, where the parties indicated that the disputed had been further narrowed, and that only 10 deposition topics remained in dispute. (See ECF 715 at 3). Accordingly, I directed the parties to meet and confer on the remaining 10 topics and, if the dispute could not be resolved, to file a letter by November 5, 2025, showing the disputed topics with color-coded highlighting to identify disputed and agreed-upon language. (*Id.*).

The parties filed their joint letter on November 5, 2025, including two different charts that used two different methods of highlighting. (See ECF 727). On November 7, 2025, I directed Plaintiffs to respond to OpenAI's chart, which contained limited argument for each disputed topic. (ECF 733). Plaintiffs filed their reply on November 12, 2025. (ECF 747).

II. LEGAL STANDARD

A. Protective Orders

District courts enjoy "broad discretion to manage the manner in which discovery proceeds." *Diamond v. 500 SLD LLC*, 21-CV-2604 (PKC), 2022 WL 956262, at *2 (S.D.N.Y. Mar. 30, 2022). "A court may issue a protective order to safeguard a party or person from annoyance, embarrassment, oppression, or undue burden or expense, but the moving party bears the burden of establishing good cause for such a protective order." *In re Lifetrade Litig.*, 17-CV-2987 (JPO) (KHP), 2023 WL 3244517, at *2 (S.D.N.Y. May 4, 2023) (internal citations and quotation marks omitted). "Ordinarily, good cause exists when a party shows that disclosure

¹ Plaintiffs suggest that "[u]sing the same double-counting [method], Defendants served over 756 topics on Plaintiffs." (ECF 688 at 1).

will result in a clearly defined, specific, and serious injury.” *Id.* (citing *In re Terrorist Attacks on Sept. 11, 2001*, 454 F. Supp. 2d 220, 222 (S.D.N.Y. 2006)). Ultimately, the discretion to grant or deny a protective order lies with the district court. *Dove v. Atl Cap. Corp.*, 963 F.2d 15, 19 (2d Cir. 1992).

B. 30(b)(6) Depositions

“All 30(b)(6) deposition topics are constrained by Rule 26(b)(3),” and must be both relevant and proportional to the needs of the case. *In re Lifetrade Litig.*, 2023 WL 3244517, at *2. The party seeking a Rule 30(b)(6) deposition must describe the deposition topics with “reasonable particularity,” meaning they must be “specific as to subject area and to have discernible boundaries.” *Id.* (citing *Winfield v. City of New York*, 15-CV-5236 (LTS) (KHP), 2018 WL 840085, at *5 (S.D.N.Y. Feb. 12, 2018)). Thus, proposed deposition topics may not be “vague or overbroad in the information they seek.” *Harger Da Silva v. New York City Transit Auth.*, 17-CV-4550 (FB) (VMS), 2023 WL 8096925, at *3 (E.D.N.Y. Nov. 20, 2023). Courts have “considerable discretion in assessing whether topics have been noticed with reasonable particularity,” and generally examine three factors: (1) the nature of the topics; (2) whether the descriptions include examples of questions and clarifying information; and (3) whether a reasonable person would understand how to prepare for the deposition. *Seliger v. Breitbart News Network, LLC*, 20-CV-2860 (ER), 2021 WL 707063, at *2 (S.D.N.Y. Feb. 22, 2021).

“When an organization, such as a corporation, partnership, or association, is issued a Rule 30(b)(6) deposition notice, the organization will designate an individual or individuals to ‘testify on its behalf about information known or reasonably available to the organization.’” *Fashion Exchange LLC v. Hybrid Promotions, LLC*, 14-CV-1254 (SHS) (OTW), 2019 WL 1533212, at *1 (S.D.N.Y. Apr. 9, 2019) (quoting Fed. R. Civ. P. 30(b)(6)). Thus, an organization “has a duty to

prepare its designated witness to answer questions to the best of the organization's knowledge." *Id.* (citing *S.E.C. v. Morelli*, 143 F.R.D. 42, 45 (S.D.N.Y. 1992)). A corporation may prepare and present multiple individuals to testify on its behalf if one person cannot address all of the topics. *In re Lifetrade Litig.*, 2023 WL 3244517, at *2. However, in order to satisfy its duty to adequately prepare a 30(b)(6) witness, a corporation's efforts need only be reasonable and performed in good faith:

The organization must make a conscientious good faith endeavor to designate the persons having knowledge of the matters identified and to prepare those persons in order that they can answer fully, completely, and unevasively, the questions posed as to the relevant subject matters. The organization must prepare the designee to the extent matters are reasonably available, whether from documents, past employees, or other sources. At the same time, a Rule 30(b)(6) witness is not expected to know and testify about every detail on a topic.

K.S. v. City of New York, 21-CV-4649 (JSR) (SLC), 24-CV-3390 (JSR) (SLC), 2025 WL 307348, at *2 (S.D.N.Y. Jan. 27, 2025) (internal citations and quotation marks omitted). "If the entity receiving the deposition notice does not possess knowledge of the matters listed in the deposition notice, then its obligations under Rule 30(b)(6) obviously cease, since the rule requires testimony only as to matters known or reasonably available to the organization." *Id.* (internal citations and quotation marks omitted). *See also Bigsby v. Barclays Capital Real Estate, Inc.*, 329 F.R.D. 78, 81 (S.D.N.Y. Jan. 25, 2019).

III. DISCUSSION

The parties' briefing indicates that the following nine topics² remain in dispute:³

- 1) At a high level, OpenAI's development of GPT-1, GPT-2, and GPT-3 and the extent to which they impacted the development of subsequent, relevant models.
- 2) Custom GPTs and ChatGPT plug-ins, including ... (E) Your knowledge of custom GPTs and plug-ins that reference news, paywalls, Journalism, News Content, or News Plaintiffs.
- 3) Your knowledge of and understanding of News Plaintiffs' Terms of Service or Terms of Use, including the provisions which govern use of News Plaintiffs' content and services.
- 4) Your awareness of whether Your Text Generation AI Models and Generative AI products and Service include and/or rely upon copyrighted content, in the past or present, including any awareness of whether Your Training Datasets include or included copyrighted content.
- 5) Your knowledge of any use of Your Generative AI Products and Services, including by third parties, to create websites or social media posts containing AI-generated content, including to promulgate misinformation, disinformation, or poor quality "pinkslime" journalism. This Topic includes websites and entities that NewsGuard has identified as "Unreliable AI-Generated News Websites." This Topic also includes Your knowledge and efforts to track and/or prevent the same.
- 6) Your knowledge of "so-called" shadow libraries such as Library Genesis ("Libgen"), Zlibrary, Bibliotik, Sci-hub, Internet Archive, Anna's Archive, Books3, or others, including the purpose of the shadow library, the content available at the shadow library, the legality of the shadow library, and the efforts or lack of efforts of the shadow library to protect the copyrights of content available at the shadow library.
- 7) Your efforts to acquire data from "so-called" shadow libraries such as Library Genesis ("Libgen"), Z-library, Bibliotik, Sci-hub, Internet Archive, Anna's Archive, Books3, or others, including, but not limited to, the identify of all individuals involved with the decision to acquire each dataset, the identity of the individuals who downloaded each dataset, the website each dataset was acquired from, the

² In OpenAI's chart and Plaintiffs' reply, the parties indicated that Topic 10 is no longer in dispute. (ECF Nos. 727-1, 747-1).

³ Disputed language is underlined.

date each dataset was first acquired, how and where each dataset was stored once acquired, and any discussion or deliberation about whether to acquire each dataset.

- 8) Your use of books datasets such as Library Genesis (“Libgen”), Z-library, Bibliotik, Sci-hub, ELibra, Internet Archive, or other in training and developing Your Text Generation AI Models including, but not limited to, the identification of which datasets were used to train each model, the identify of all individuals involved in the decision to use the datasets to train each model, how many copies of each dataset were made in training and development of each model, any modifications or analyses conducted on the datasets, and any discussion or deliberation about whether to use each dataset to train each model.
- 9) The storage, organization, preservation, and/or deletion of data or content used for any stage of training and/or developing Your Text Generative AI Models and Generative AI Products and Services. This Topic includes where and how You stored the data or content, any modifications to the data or content and the individuals involved in storing the data or content.

(See ECF 747-1).

OpenAI appears to challenge Topics 2 and 5 on the grounds of relevance. (See ECF 747-1 at 3-4, 6-7). Subpart (e) of Topic 2 and Topic 5 are clearly relevant to Plaintiffs’ contributory infringement claims and the fourth fair use factor.

Next, OpenAI’s contends for the majority of the deposition topics listed above (Topics 2-6) that the “knowledge” component is not proportional to the needs of the case because it would require OpenAI to interview every employee at the company to properly prepare their 30(b)(6) witness. (See ECF 747-1). That is not so. Rule 30(b)(6) only requires a corporation to prepare their 30(b)(6) designees to the extent that matters are known or “reasonably available.” *K.S.*, 2025 WL 307348, at *2. This only requires OpenAI to conduct a reasonable investigation—not exhaustive interviews of every single OpenAI employee. If, after a reasonable investigation, OpenAI determines that it does not have institutional knowledge on the particular topic or question, it can testify as to that fact and the lack of knowledge will be

imputed to the corporation. See *Fashion Exchange LLC v. Hybrid Promotions, LLC*, 333 F.R.D. 302, 304 (S.D.N.Y. 2019) (“[T]he Court warned both parties that [30(b)(6)] depositions require advance preparation of the witnesses on ‘knowledge that can be imputed to the entire organization, not just their personal knowledge.’”); *Fashion Exchange*, 2019 WL 1533212, at *1 (“As a result, the organization has a duty to prepare its designated witness to answer questions to the best of the organization’s knowledge.”) (emphasis added).

Next, OpenAI asserts that because Topics 6, 7, and 8 contain non-exhaustive lists of shadow libraries, these topics are vague and lack the requisite particularity under Rule 30(b)(6). (See, e.g., ECF 747-1 at 7 (“The topic lacks the requisite particularity to the extent it purports to cover all ““so-called” shadow libraries.””). OpenAI cites no authority to suggest that 30(b)(6) deposition topics may only include finite lists. Indeed, weighing the three factors set out in *Seliger*, because the nature of the topics is not complicated and includes examples of the types of shadow libraries that Plaintiffs wish to discuss, a reasonable person would be able to understand how to prepare a 30(b)(6) designee to testify on this topic. 2021 WL 707063, at *2. OpenAI’s concerns are also belied by the fact that it only needs to conduct a reasonable investigation—this does not require OpenAI to discern every possible website that could possibly fit into the category of “shadow libraries” to adequately prepare its designee(s).

The parties also dispute the inclusion of the phrase “high-level” in Topic 1. OpenAI asserts that GPT-1, GPT-2, and GPT-3 are not in the scope of the case, and thus only wishes to provide “high-level” testimony regarding these models and how they impacted later model development, while News Plaintiffs contend that those models are in the case. (ECF 747-1 at 2). Judge Stein’s recent Opinion & Order on OpenAI’s motion to strike portions of the Consolidated Class Action Complaint clearly states that GPT-3 is at issue. (See ECF 707 at 1) (“[T]he Court

limited the models to be included in the Consolidated Class Action Complaint to GPT-3, GPT-3.5...” (emphasis added). Moreover, News Plaintiffs are entitled to ask questions at deposition on any of OpenAI’s models that are listed in their operative complaint(s). (*See, e.g.*, ECF 72 ¶ 85) (“The training dataset for GPT-2 includes an internal corpus OpenAI built called ‘WebText,’ which includes ... a staggering amount of scraped content from The Times.”). Including the modifier “high-level” is thus an inappropriate limitation. If OpenAI believes that certain questions are overbroad or not in the scope of the case, they can meet and confer with Plaintiffs and make their objections on the record.

With respect to the remaining disputes, because none of the topics appear to explicitly request or implicate privileged communications, OpenAI should raise any privilege objections during the deposition, and the Court can revisit the issue at a later time if necessary. *DDK Hotels, LLC v. Williams-Sonoma, Inc.*, 19-CV-226 (ILG) (CLP), 2022 WL 2702378, at *15 (E.D.N.Y. Feb. 11, 2022). Nor is Topic 4 a backdoor for OpenAI’s contentions—this topic does not require OpenAI to identify the bases for any of its claims, it simply seeks testimony on OpenAI’s state of mind, which is relevant to willfulness and OpenAI’s good faith/innocence defense. To the extent that OpenAI claims that any particular topic is duplicative of prior 30(b)(6) testimony, (*see, e.g.*, ECF 747-1 at 10) (explaining that Topic 9 “seeks duplicative 30(b)(6) testimony”), Plaintiffs are limited to questions concerning new facts and information obtained since prior depositions.

IV. CONCLUSION

Accordingly, for the foregoing reasons, OpenAI’s request for a protective order is **DENIED**. The language “high-level” should be stricken from Topic 1, and the remaining disputed language should be retained for Topics 2 – 9.

The Clerk of Court is respectfully directed to close ECF Nos. 727 and 747.

SO ORDERED.

Dated: November 24, 2025
New York, New York

s/ Ona T. Wang
Ona T. Wang
United States Magistrate Judge