



May 23, 2024

Hon. Sidney H. Stein  
Daniel Patrick Moynihan  
United States Courthouse  
500 Pearl St.  
New York, NY 10007-1312

cc: All Counsel of Record Line (via ECF)

**Re: New York Times Company v. Microsoft Corp., et al., Case No 1:23-cv-11195-SHS**

Dear Judge Stein:

Pursuant to Rule 2(G) of Your Honor’s Individual Practices, OpenAI seeks an informal discovery conference concerning two issues. *First*, Plaintiff, The Times, refuses to commit to a substantial completion deadline in response to OpenAI’s requests for production, served on March 8, despite the Court’s adoption of Plaintiff’s own expedited discovery schedule. OpenAI seeks an order compelling Plaintiff to substantially complete document production by a date certain. *Second*, Plaintiff has refused to produce certain documents underlying core allegations of Plaintiff’s complaint relating to allegedly infringing outputs from ChatGPT. Plaintiff’s privilege and work product claims are baseless, and OpenAI seeks an order compelling production.

Substantial Production Deadline. The parties submitted competing proposals for the case schedule in the Rule 26(f) report on March 8. Dkt. 72 at 14-15. Plaintiff sought an accelerated discovery schedule, requiring substantial production by June 14 for RFPs served by February 28, and close of fact discovery on September 17. *Id.* OpenAI sought a longer schedule, with substantial production by February 7, 2025 and close of fact discovery on April 7, 2025. *Id.* OpenAI served RFPs on Plaintiff on March 8, the same day that the Rule 26(f) Report was filed. On May 3, 2024, the Court accepted Plaintiff’s schedule. Dkt. 112.

In light of the short discovery period, OpenAI asked Plaintiff to commit to a substantial production deadline of June 24 for the March 8 RFPs. This mirrors the amount of time the Court ordered for OpenAI to substantially complete production in response to Plaintiff’s February 23 RFPs. After meeting and conferring, Plaintiff refused to commit to this deadline, or any substantial completion deadline at all. We respectfully request the court to compel Plaintiff to make a substantial production of documents responsive to the March 8 RFPs by June 24.

Plaintiff’s Regurgitation Efforts. The centerpiece of Plaintiff’s complaint is a claim that OpenAI’s ChatGPT large language models (LLMs) “output near-verbatim copies of significant portions of Times Works when prompted to do so.” Compl. ¶ 98. That assertion is purportedly supported by a lengthy exhibit—Exhibit J—containing one hundred allegedly infringing outputs generated using ChatGPT. *See id.* Ex. J. Plaintiff asserted this evidence shows that “individuals . . . will likely” access Times’s content “without having to pay for it” to such a degree that it will “divert readers, including current and potential subscribers, away from The Times, thereby reducing” The Times’s revenues. Compl. ¶ 157. Thus, a central question in this case is exactly

how Plaintiff was able to generate these outputs, which appear to be the fruits of Plaintiff's prolonged and extensive efforts to manipulate the ChatGPT LLMs.

Accordingly, before discovery even began, OpenAI put Plaintiff on notice about its obligation to preserve materials related to the creation of Exhibit J and began requesting information about it. Ex. A (Feb. 9 Letter). Then, on March 8, OpenAI served Requests for Production seeking all such materials, including (a) documents to show OpenAI accounts used by Plaintiff and its agents, including those used to generate the outputs cited in or referred to in the complaint (RFP Nos. 7 & 21); (b) documents and communications regarding attempts—including failed attempts—to reproduce the published works, including those “relat[ed] to the creation of Exhibit J” (RFP Nos. 2 & 20); and (c) documents sufficient to show the process for obtaining the output cited or referred to in the complaint (RFP No. 23). *See* Ex. B (Requests for Production).

On April 8, Plaintiff objected, in part, on the basis of the attorney-client privilege and work-product doctrine, but does not dispute that the materials sought (other than the requested account information) are relevant and proportional to the needs of the case. *See* Ex. C (Plaintiff's Objections); *see also* Ex. D (May 14 Email). After meeting and conferring, however, Plaintiff indicated that it nevertheless intends to withhold “outputs that weren't cited in the complaint.” Ex. E at 3 (May 22 Email); *see also* Ex. H (May 17 Email); Ex. G (May 15 Email); Ex. F at 1-2 (May 7 Letter). On other related categories, Plaintiff refuses to provide a clear answer. *See* Ex. H (May 17 Email). Additionally, Plaintiff refuses to produce *any* documents sufficient to identify their OpenAI accounts, including those involved in the creation of outputs referenced in the Complaint. *Id.* For the reasons below, Plaintiff should be compelled to provide responsive documents.

The information requested is relevant, and indeed, for all the materials requested except for OpenAI account information, Plaintiff has agreed that it will not withhold those materials on relevance grounds. Plaintiff implicitly recognizes that the complaint has put directly at issue whether ChatGPT outputs like those generated by Plaintiff will actually “divert readers” and siphon revenue from The Times, which depends in part on how difficult it was for Plaintiff to generate those outputs and whether its methodology accurately approximates realistic use of ChatGPT LLMs. Compl. ¶ 157. As to the requested OpenAI account information, that is relevant because it would potentially enable OpenAI to fill gaps in the information provided. For instance, the Complaint states that there are numerous parameters one must use to “tune” the model in order to create a desired output, like “model,” “temperature,” “maximum length,” “top p,” “frequency penalty,” and “presence penalty.” *E.g.*, Compl. ¶ 140. Plaintiff should have preserved that information and should produce it. But if it did not, OpenAI might be able to ascertain it from its own information using the identity of the relevant account; and even if Plaintiff does provide that information, the account information might help OpenAI verify its accuracy.

There is no burden to produce this information. Based on the parties' exchanges, Plaintiff clearly possesses an easily identifiable set of materials responsive to these requests. Plaintiff has represented that the ChatGPT outputs identified in its complaint were procured by a single independent researcher that Plaintiff retained for this litigation. Ex. A (Feb. 9 Letter). That lone researcher presumably maintained basic records documenting his or her efforts.

Plaintiff's claim for work-product protection over OpenAI account information and prompt and output data has been waived because that data and information was voluntarily disclosed to

OpenAI in the course of interacting with ChatGPT LLMs. *See, e.g., New York Times Co. v. United States Dep't of Just.*, 939 F.3d 479, 494 (2d Cir. 2019) (explaining that “[d]isclosing work product to [an] adversary” waives work-product protection); *see also In re Steinhardt Partners, L.P.*, 9 F.3d 230, 235 (2d Cir. 1993).<sup>1</sup> (Any work-product protection over OpenAI user account, prompt and output data was also waived for the separate, additional reason that Plaintiff put that material “at issue,” as discussed below.)

With respect to related materials that Plaintiff may be withholding (such as analysis or data regarding the failure of attempts to use ChatGPT LLMs to generate infringing outputs), both work-product protection and attorney-client privilege were waived by Plaintiff’s decision to put those materials at issue in this case. Such a waiver can be found when, as here, “a party advances a claim to a court . . . while relying on its privilege to withhold from a litigation adversary materials that the adversary might need to effectively contest or impeach the claim.” *New York Times Co.*, 939 F.3d at 495. Similarly, attorney-client privilege can be waived “if a party puts the privileged communication at issue by relying on it to support a claim or defense.” *Financial Guaranty Insurance Co. v. Putnam Advisory Company, LLC*, 314 F.R.D. 85, 90 (S.D.N.Y. 2016).

That describes precisely the situation here. For example, Plaintiff’s presentation creates the potentially false impression that such regurgitation is easy for the typical user to reproduce. But if generating the outputs displayed in the complaint took significant effort or trial and error, then the behavior highlighted in the complaint is not probative of typical ChatGPT behavior. In fact, it would suggest that the threat of any harm via reduced readership is exaggerated. In short, OpenAI should be permitted to obtain from Plaintiff “materials that [it] might need to effectively contest or impeach [the Plaintiff’s] claim.” *New York Times Co.*, 939 F.3d at 495.

Analogous cases in this District support that conclusion. In *In re Commodity Exch., Inc., Gold Futures & Options Trading Litig.*, the plaintiffs’ complaint included allegations of statistical correlations that supported the plaintiffs’ claims. 2019 WL 13046984, at \*1-2 (S.D.N.Y. Feb. 25, 2019). The Court rejected plaintiffs’ assertions of work-product protection and ordered disclosure of reports, analysis, and “any materials underlying those reports and analysis” that were connected to the statistical results in the complaint, concluding that “the withholding of information that would tend to undermine key statistical conclusions [in the] complaint would . . . result in a selective and misleading presentation of evidence to the disadvantage of the adversary.” *Id.* at \*2-3 (cleaned up); *see also Financial Guaranty Ins. Co. v. Putnam Advisory Co.*, 314 F.R.D. 85, 89-90 (S.D.N.Y. 2016) (including certain assertions in the complaint regarding an economic analysis along with quotations from that analysis triggered waiver of attorney client privilege and work product protection over the entire underlying analysis because the company “put[] the privileged communication at issue by relying on it”).

Accordingly, OpenAI respectfully requests that the Court compel Plaintiff to produce documents responsive to the RFPs at issue.

---

<sup>1</sup> There is no plausible claim for attorney-client privilege for this information, which was directly given to OpenAI.

Respectfully,

KEKER, VAN NEST &  
PETERS LLP

/s/ Michelle Ybarra  
Michelle Ybarra \*

LATHAM & WATKINS  
LLP

/s/ Joseph R. Wetzel  
Joseph R. Wetzel

MORRISON &  
FOERSTER LLP

/s/ Allyson R. Bennett  
Allyson R. Bennett \*

\* The parties use electronic signatures with consent in accordance with Rule 8.5(b) of the Court's ECF Rules.