January 5, 2026

Hon. Ona T. Wang
United States Magistrate Judge
Southern District of New York

Re: *In re: OpenAI, Inc., Copyright Infringement Litigation*, 1:25-md-03143 (SHS) (OTW); this document relates to No. 1:23-cv-11195 (SHS) (OTW)

Dear Magistrate Judge Wang:

Plaintiffs The New York Times Company, Daily News, LP, et al., The Intercept, the Center for Investigative Reporting, and the *Ziff Davis* Plaintiffs (collectively, "News Plaintiffs") respectfully submit that OpenAI has engaged in sanctionable conduct in failing to preserve its output logs, including in a manner that violates both the letter and spirit of this Court's preservation and production orders. Dkt. 33; Dkt. 734. Specifically, after this Court ordered OpenAI to produce 20 million logs over OpenAI's vociferous and repeated objections, OpenAI substituted millions of conversations that it was ordered to produce with other conversations – seemingly because it had deleted millions of the selected logs. OpenAI also obstructed News Plaintiffs' review of the logs it did produce by improperly redacting conversations and altering the conversation IDs so that they cannot be matched with the original list shown to News Plaintiffs. OpenAI has refused to answer News Plaintiffs' questions about the deleted and substituted logs. News Plaintiffs therefore request this Court to: (1) order OpenAI to show cause why it should not be held in contempt for violating this Court's orders, and (2) set an evidentiary hearing following the close of expert discovery to determine appropriate remedies or sanctions.

A. OpenAI's Destruction of the Output Logs and Failure to Comply with its Court-Ordered Discovery Obligations

This Court is familiar with News Plaintiffs' longstanding discovery efforts into OpenAI's output logs, and News Plaintiffs incorporate by reference the numerous filings on this issue. *See* Appendix A. From the inception of this litigation, OpenAI has ignored its preservation obligations and obstructed News Plaintiffs' access to critical evidence. On at least 24 separate occasions since the filing of The Times's complaint, OpenAI compressed tens of billions of consumer ChatGPT logs, which rendered them unsearchable and therefore functionally unavailable. Dkt. 935 at 4, n.4. OpenAI further: (a) destroyed virtually all of its API logs; (b) represented that it does not have possession, custody, or control of the enterprise ChatGPT logs; and (c) destroyed billions of consumer ChatGPT logs where the user turned ChatGPT's preservation function off or were the subject of a user-initiated deletion.

At the request of News Plaintiffs, and following extensive motion practice, this Court entered two orders relating to the output log data: (1) the May 13 Order directing OpenAI to "preserve and segregate all output log data that would otherwise be deleted on a going forward basis until further order of the Court," Dkt. 33 (the "Preservation Order"), and (2) the November 7 Order directing OpenAI to produce the 20 million deidentified consumer ChatGPT output logs, Dkt. 734 (the "Production Order").

Within weeks of this Court's Preservation Order, OpenAI began publicly criticizing its court-ordered preservation obligations. *See, e.g.*, Ex. E (OpenAI June 5, 2025 blog post titled "How we're responding to The New York Times' data demands in order to protect user privacy"); Ex. F (June 5, 2025 Sam Altman tweet attacking Preservation Order). While purporting to be working to produce the output logs, OpenAI was in fact deleting millions of logs in the sample selected by News Plaintiffs. In a letter dated September 12, 2025, OpenAI advised News Plaintiffs that it had to replace "approximately 5%" of the logs News Plaintiffs had originally requested because they were no longer in OpenAI's possession (*i.e.*, OpenAI deleted them). Ex. A. In October 2025, News Plaintiffs learned that some of these deletions occurred during the pendency of the Preservation Order. Ex. B.

OpenAI first made 20 million de-identified consumer ChatGPT logs available to News Plaintiffs on Monday, December 15, although News Plaintiffs believe that two million of these – 10%, not the 5% previously indicated – were not the logs that News Plaintiffs had selected for sampling and OpenAI had been ordered to produce. Figuring out which logs were missing from the sample was complicated because OpenAI "hashed" the conversation IDs of all the 20 million logs. However, when our experts compared the time stamps, we learned that about 10% of the timestamps do not match what News Plaintiffs had requested on August 11 (and what OpenAI had agreed, and was ordered, to produce). Ex. C at 2.

In the parties' correspondence and during a meet and confer on December 23, OpenAI admitted that: (i) at least 5% of the logs were replaced because OpenAI had deleted the data, and (ii) OpenAI "filtered" logs that were associated with OpenAI, Microsoft, or government accounts without informing News Plaintiffs or seeking their consent.[1] On that meet and confer, OpenAI agreed to "look into" providing metadata so that News Plaintiffs may identify which logs were replaced (which it has still not produced), but refused to provide information indicating the reason(s) for each replacement.

That is not all OpenAI has done to obstruct News Plaintiffs. OpenAI made grossly overbroad and inappropriate redactions, including instances where OpenAI redacted URLs corresponding to ***News Plaintiffs' own websites*** and articles and to conversations in which users requested content from news publishers. Ex. C at 13. While OpenAI said it is "willing to work with News Plaintiffs" about these plainly improper redactions, OpenAI declined to explain how the redactions were applied or to remove all such redactions. Ex. C at 3-4. News Plaintiffs' investigation is ongoing, but it appears that OpenAI has applied over 19 billion redactions to the consumer ChatGPT logs, including to URLs (over 1.1 million), publication names, city names, headlines, and bylines of articles, among other information. Ex. D.

News Plaintiffs and their experts have spent a substantial amount of time trying to remedy these deficiencies. As part of these efforts, they requested that OpenAI identify the dates on which it had deleted the approximately two million conversations. Ex. C at 2. This is critically important for at least two reasons. *First*, as OpenAI instituted Plaintiff-specific blocks after these cases were

[1] Some of the logs were replaced because they were associated with GPT models post-dating GPT-4o mini. Such replacements are appropriate, but OpenAI has refused to identify which logs were replaced on that basis.

filed, earlier conversations are likely to contain a higher percentage of relevant outputs. *Second*, Sam Altman and OpenAI chose to widely publicize to OpenAI's users that their conversations could be disclosed to News Plaintiffs in this litigation, which may well have caused some users to ask OpenAI to delete conversations in which they called for copyrighted content from News Plaintiffs. OpenAI has refused to provide these dates. Ex. C at 1.

B. An Evidentiary Hearing Is Appropriate to Determine Remedies and Sanctions Arising From OpenAI's Litigation Misconduct

OpenAI's failure to preserve relevant evidence and efforts to obstruct News Plaintiffs' review of that evidence likely constitutes sanctionable conduct. *Ransom v. Andrews*, No. 21-cv-6343, 2022 WL 16555362, at *3 (S.D.N.Y. Oct. 31, 2022); *Zubulake*, 220 F.R.D. 212, 218 (S.D.N.Y. 2003). Once the obligation to preserve relevant electronically stored information ("ESI") attaches, the party in possession of that ESI must take "reasonable steps" to preserve it. Fed. R. Civ. P. 37(e); *Charlestown Cap. Advisors, LLC v. Acero Junction, Inc.*, 337 F.R.D. 47, 61 (S.D.N.Y. 2020); *Lokai Holdings LLC v. Twin Tiger USA LLC*, 2018 WL 1512055, at *11 (S.D.N.Y. Mar. 12, 2018). The "reasonable steps" inquiry has been equated to "roughly a negligence standard." *Leidig v. Buzzfeed, Inc.*, 2017 WL. 6512353, at *10 (S.D.N.Y. Dec. 19, 2017)."Once the duty to preserve attaches, any destruction of documents is, *at a minimum*, negligent." *Zubulake*, 20 F.R.D. at 220 (emphasis added). Courts have broad discretion to fashion appropriate corrective measures to address discovery violations. *Eur. v. Equinox Holdings, Inc.*, 592 F. Supp. 3d 167, 175 (S.D.N.Y. 2022).

OpenAI repeatedly disregarded its obligation to take "reasonable steps" to preserve critically important output log data "that should have been preserved in the anticipation or conduct of litigation." Fed. R. Civ. P. 37(e). News Plaintiffs have spent thousands of hours pursuing discovery into OpenAI's output logs, and this Court has addressed dozens of motions, held three separate confidential conferences, and issued the Preservation Order and Production Order. OpenAI violated both those orders by failing to produce 5-10% of the small sample of logs News Plaintiffs requested because it deleted them or "filtered" them out — and heavily redacted the remainder in a manner that impedes News Plaintiffs' analysis. The methodology by which OpenAI (over)redacted the logs, the identification of the replaced logs and the reasons for the replacements, and the timing of the deletions of the logs (including whether those deletions occurred during OpenAI's public outcries), all bear on any bias or other issues in searching and extrapolating the sample.

News Plaintiffs respectfully request this Court: (1) issue an order for OpenAI to show cause why it should not be held in contempt for violating this Court's orders, and (2) set an evidentiary hearing following the close of expert discovery to determine appropriate remedies or sanctions from OpenAI's conduct. This letter highlights only some of OpenAI's discovery misconduct relating to output logs, and is not intended to fully address prejudice from both the discovery failures listed here and those that have already been raised before this Court. *See* Appendix A.

Dated January 5, 2026

Respectfully submitted,

/s/ Steven Lieberman

Steven Lieberman
Rothwell, Figg, Ernst & Manbeck, P.C.
Counsel for The New York Times and Daily News Plaintiffs

/s/ Davida Brook

Davida Brook
Susman Godfrey L.L.P.
Counsel for The New York Times and News Plaintiffs' Liaison Counsel

APPENDIX A: SELECTED INDEX OF COURT FILINGS RELATED TO OPENAI'S OUTPUT LOGS

1. January 13, 2025: News Plaintiffs' letter motion requesting this Court order OpenAI to: "(1) preserve its output log data, including output log data that contains News Plaintiffs' content, hereinafter, and (2) identify all output log data that has been destroyed, when it was destroyed, and whether it is retrievable." *Times* Dkt. 379.

2. January 16, 2025: OpenAI filed its opposition to News Plaintiffs' letter motion (*Times* Dkt. 379). *Times* Dkt. 421.

3. January 24, 2025: The Court denied News Plaintiff's letter motion (*Times* Dkt. 379), instead directing the parties "to meet and confer regarding the privacy and technological considerations implicated in Plaintiffs' request and file a joint status letter on the docket by February 13, 2025, regarding the status of this issue." *Times* Dkt. 441 (emphasis omitted).

4. February 13, 2025: The parties submitted a joint letter with a chart summarizing the status of various discovery disputes that were addressed at the January 22, 2025 conference. *Times* Dkt. 462. News Plaintiffs' noted that " OpenAI has refused to give The Times and Daily News Plaintiffs a clear answer as to what amount of output log data OpenAI deleted for each product at issue, and OpenAI will still not agree to preserve relevant output log data." *Id.*, p. 3. News Plaintiffs' renewed their request from their earlier motion (*Times* Dkt. 379), for this Court order OpenAI to: "(1) preserve its output log data, including output log data that contains News Plaintiffs' content, hereinafter, and (2) identify all output log data that has been destroyed, when it was destroyed, and whether it is retrievable." *Id.*

5. March 25, 2025: News Plaintiffs' letter motion requesting that this Court order OpenAI to: "provide in writing no later than March 31, 2025: (1) an identification of all output log data that has been destroyed, including an estimated volume thereof based on the best information available to OpenAI, since December 27, 2023, when The Times filed its original Complaint, (2) when such data was destroyed, and (3) whether such data is retrievable." *Times* Dkt. 483 at 2.

6. March 26, 2025: This Court issued an Order directing OpenAI to file a response to News Plaintiffs' March 25, 2025 letter (*Times* Dkt. 483) by March 31, 2025. *Times* Dkt. 484.

7. March 31, 2025: OpenAI filed its opposition to News Plaintiffs' letter motion (*Times* Dkt. 483). *Times* Dkt. 491.

8. April 30, 2025: This Court issued Case Management Order No. 1, which noted that "[i]f there is currently a dispute about the scope or degree of preservation of

output log data, the parties should identify the dispute in the Report. This Order does not resolve or address any such dispute." *MDL* Dkt. 14 at 4, n.1.

9. May 7, 2025: News Plaintiffs' filed leave to file a Supplemental Letter regarding News Plaintiffs' then-pending Motion to Compel OpenAI to identify the output logs it has destroyed since The Times filed its original Complaint on December 27, 2023. *MDL* Dkt. 21.

10. May 9, 2025: OpenAI filed a response to News Plaintiffs' request to file a Supplemental Letter (*MDL* Dkt. 21). *MDL* Dkt. 24.

11. May 9, 2025: The Court granted News Plaintiffs request to file a Supplemental Letter. *MDL* Dkt. 26.

12. May 12, 2025: News Plaintiffs filed a Supplemental Letter to this Court attaching OpenAI's March 31, 2025 letter, where OpenAI admitted it destroyed over 8 billion ChatGPT Free, Pro, and Plus output logs since December 27, 2023. The letter did not identify the volume of output log data OpenAI had destroyed (or retained) for the ChatGPT Enterprise or API products. Dkt. 29; Dkt. 29-2 (March 31 letter).

13. May 13, 2025: The Court scheduled an In-Person Status Conference regarding potential spoliation motions for the deletion of output log data on May 27, 2025. *MDL* Dkt. 32. The Court also ordered supplemental briefing, directing the News Plaintiffs' to file a supplemental letter by May 16, 2025, and OpenAI to file a supplemental opposition by May 23, 2025. *Id.*

14. May 13, 2025: This Court issued a Preservation Order directing OpenAI to "preserve and segregate all output log data that would otherwise be deleted on a going forward basis until further order of the Court." *MDL* Dkt. 33.

15. May 15, 2025: This Court denied OpenAI's motion for reconsideration. *MDL* Dkt. 42. The Court permitted OpenAI to renew its motion with their supplemental briefing set for May 23, 2025. *Id.*

16. May 15, 2025: OpenAI filed a motion for reconsideration of the May 13, 2025, Preservation Order. *MDL* Dkt. 40.

17. May 16, 2025: News Plaintiffs filed a supplemental brief requesting: (1) that the interim preservation order remain in place, (2) for OpenAI to provide a declaration and relevant documents, and make available for deposition one or more knowledgeable 30(b)(6) witnesses, on topics concerning the destroyed output log data, and (3) for OpenAI to search over the full extant output log data for News Plaintiffs' content. *MDL* Dkt. 44 - sealed.

18. May 23, 2025: OpenAI filed a supplemental opposition. *MDL* Dkt. 65.

19. May 27, 2025: This Court held a status conference "regarding potential spoliation motions for the deletion of output log data." (*MDL* Dkt. 32). At the conference, this Court excluded Enterprise ChatGPT output log data from the Preservation Order but maintained the Preservation Order with respect to the consumer ChatGPT and API output log data. *MDL* Dkt 291 (Transcript of May 27, 2025 Hearing, Afternoon Session), p. 52.

20. May 29, 2025: This Court issued an order directing the parties to begin engaging in sampling of the 30-day tables of consumer ChatGPT output log data, and denying without prejudice News Plaintiffs' request for additional discovery or an additional 30(b)(6) deposition on the preservation of the output log data. *MDL* Dkt. 79. The order directed OpenAI to submit a proposal for the sampling by June 6, 2025. The order also denied OpenAI's motion for reconsideration of the Court's May 13, 2025, Preservation Order. *Id.*

21. June 4, 2025: OpenAI filed a Rule 72 Objection to this Court's May 13, 2025, Preservation Order. *MDL* Dkt. 92.

22. June 6, 2025: OpenAI submitted its proposal for sampling. *MDL* Dkt. 104. Because OpenAI did not share the proposal with News Plaintiffs, or seek their input on the proposal, News Plaintiffs requested leave to submit a response to the proposal by June 11, 2025. *MDL* Dkt. 98.

23. June 10, 2025: OpenAI filed a Letter Motion requesting that the Court modify its May 13, 2025, Preservation Order. *MDL* Dkt. 127.

24. June 11, 2025: This Court granted News Plaintiffs' request to respond to OpenAI's sampling proposal, directing News Plaintiffs to submit their response by June 13, 2025. *MDL* Dkt. 161.

25. June 13, 2025: News Plaintiffs filed their opposition to OpenAI's request to modify the May 13, 2025, Preservation Order. *MDL* Dkt. 169. News Plaintiffs also filed their response to OpenAI's proposed sampling order. *MDL* Dkt. 171.

26. June 18, 2025: This Court ordered the parties to meet and confer on their respective proposed sampling orders. *MDL* Dkt. 217. The order also scheduled a confidential settlement conference for June 26, 2025. *Id.*

27. June 25, 2025: The parties submitted a joint letter on a proposed sampling protocol. *MDL* Dkt. 263. The letter set forth respective positions for News Plaintiffs and OpenAI. OpenAI also submitted its own proposed order for a sampling protocol. *MDL* Dkt. 268.

28. June 26, 2025: This Court held a confidential settlement conference.

29. June 26, 2025: The Court, for the reasons stated on the record, denied OpenAI's objection to the Court's May 13, 2025, Preservation Order. *MDL* Dkt. 271.

30. July 30, 2025: News Plaintiffs and OpenAI filed cross motions with respect to sampling of the 30-day tables of consumer ChatGPT output log data. *MDL* Dkt. 398; *MDL* Dkt. 404.

31. July 30, 2025: News Plaintiffs filed a letter motion to compel OpenAI to produce 120 million total samples (5 million per month) of output log data from December 2022 through November 2024. *MDL* Dkt. 394.

32. August 15, 2025: This Court held a second confidential settlement conference.

33. August 19, 2025: This Court entered an order setting forth the sampling protocol for the 30-day tables of consumer ChatGPT output log data. *MDL* Dkt. 463.

34. September 20, 2025: The Parties filed a joint letter regarding the sampling analysis of the preserved tables of consumer ChatGPT output log data. *MDL* Dkt. 597.

35. September 26, 2025: This Court held a third confidential settlement conference.

36. October 9, 2025: The parties filed a stipulation and order to terminate OpenAI's ongoing obligations under the Preservation Order at ECF 33. *MDL* Dkt. 629.

37. October 15, 2025: News Plaintiffs filed a letter motion to compel OpenAI to immediately produce the agreed-upon 20 million records of consumer ChatGPT output log data. *MDL* Dkt. 656.

38. October 30, 2025: The parties filed supplemental briefs regarding News Plaintiffs' motion to compel OpenAI to produce the 20 million records of consumer ChatGPT output log data. *MDL* Dkt. 717; *MDL* Dkt. 719.

39. November 7, 2025: This Court ordered OpenAI to produce the 20 million consumer ChatGPT output logs. *MDL* Dkt. 734.

40. November 12, 2025: OpenAI filed a letter requesting a stay and reconsideration of this Court's November 7 Order. *MDL* Dkt. 742.

41. November 13, 2025: This Court denied OpenAI's request for a stay of the November 7 Order and ordered OpenAI to file a reply in support of its request for reconsideration. *MDL* Dkt. 750.

42. November 14, 2025: OpenAI filed a reply brief in support of its request for reconsideration. *MDL* Dkt. 752.

43. November 24, 2025: OpenAI filed a Rule 72 Objection to this Court's November 7 Order. *MDL* Dkt. 840.

44. December 2, 2025: This Court issued an opinion and order denying OpenAI's request for reconsideration and directing OpenAI to produce the 20 million consumer ChatGPT logs. *MDL* Dkt. 896.

45. December 8, 2025: OpenAI filed a request to stay the December 2 Order. *MDL* Dkt. 909.

46. December 8, 2025: This Court issued a third order denying OpenAI's request to stay the December 2 Order. *MDL* Dkt. 932.

47. December 10, 2025: OpenAI filed an updated Rule 72 Objection to this Court's November 7 and December 2 Orders. *MDL* Dkt. 934.

48. December 19, 2025: News Plaintiffs filed their opposition to OpenAI's updated Rule 27 Objection to this Court's November 7 and December 2 Orders. *MDL* Dkt. 987.

49. December 23, 2025: OpenAI filed a Letter Motion for leave to file a reply brief in support of its Rule 72 Objection to this Court's November 7 and December 2 Orders. *MDL* Dkt. 1004.

50. January 5, 2025: This Court denied OpenAI's objections to the November 7 and December 2 Orders. *MDL* Dkt. 1021.