

SUSMAN GODFREY L.L.P

June 3, 2024

Hon. Sidney H. Stein
United States District Judge
Southern District of New York
500 Pearl Street, Courtroom 23A
New York, New York 10007

Re: *The New York Times Company v. Microsoft Corporation, et al.*,
Case No.: 23-cv-11195-SHS: Discovery Dispute Regarding RFPs

Dear Judge Stein:

Plaintiff The New York Times Company (“The Times”) requests a conference for three disputes with OpenAI relating to The Times’s First Set of Requests for Production (“RFPs”).¹

1. OpenAI Is Improperly Limiting Discovery to Specific Large Language Models (RFP Nos. 1, 3, 6, 8, 14).

OpenAI refuses to provide discovery regarding any of the work it did to create large language models (“LLMs”) developed before 2022 – specifically, GPT (released in 2018), GPT-2 (released in 2019), and GPT-3 (released in 2020), each of which The Times alleges was unlawfully trained with Times content. Compl. ¶¶ 83-92. Yet OpenAI is unilaterally limiting discovery to two later-in-time models—GPT-3.5 (released in 2022) and GPT-4 (released in 2023)—which it claims are the only models used for ChatGPT, the user-facing chatbot released in November 2022. Ex. B at 5. This limitation affects RFPs 1, 3, 6, 8, and 14, which appropriately seek information about all of Defendants’ infringing models—not just their newer ones. *See* Ex. A. OpenAI has conceded the relevance of these requests by agreeing to produce documents in response to each, *see* Ex. B, but has carved out any documents that relate to its earlier-in-time models.

OpenAI’s statute-of-limitations argument against the earlier models (raised in its partial motion to dismiss, Dkt. 52 at 15) provides no basis to withhold discovery. OpenAI never moved to stay discovery (or any aspect thereof), and it would be inappropriate to unilaterally impose such a stay.² But even if that portion of its motion to dismiss were granted, discovery into the earlier models would still be warranted because the development of the earlier models is relevant to the development of the later models (GPT-3.5 and GPT-4).

While OpenAI asserts that each model is “trained from scratch on a new dataset” (Ex. E at 2), publicly available information suggests otherwise. OpenAI’s own blog posts show that GPT-3.5 was derived from GPT-3, just as its name suggests. In early 2022, OpenAI created an updated

¹ The parties met and conferred by videoconference on May 6, and exchanged letters and emails before and after. The Times’s RFPs are attached as Exhibit A, OpenAI’s Responses are attached as Exhibit B, and the parties’ written correspondence is attached as Exhibits C, D, and E.

² Any such motion would likely have been denied. There is a “general presumption that a motion to dismiss does not stay discovery.” *Bertrand v. Dep’t of Educ., Archdiocese of N.Y.*, 2023 WL 2776015, at *3 (S.D.N.Y. Apr. 4, 2023).

June 3, 2024

Page 2

version of GPT-3 called “InstructGPT” that it admitted – in a blog post that it seems to have removed – was used to train GPT-3.5.³

Moreover, even if each model were “trained from scratch,” OpenAI cannot dispute that it leveraged its work on prior models when developing new versions, including identifying what training processes and mixes of training data worked well. In a June 2020 blog post discussing GPT-3, OpenAI explained that it was releasing an interface that would allow developers to interact directly with its models “to see what challenges arise when AI systems are used in the real world” and to “guide our efforts to understand how deploying future AI systems will go.”⁴ The Times is entitled to know precisely how OpenAI’s learnings from the development of earlier models impacted the development of later models, particularly as it pertains to the use of copyrighted publisher content.

Finally, even if the later models were developed in a vacuum without any reliance on or reference to the work done on the prior models (which is not true), The Times would still be entitled to this discovery. OpenAI once again misconstrues The Times’s claims as limited to the outputs generated by ChatGPT, which they are not. Ex. D at 2; Ex. E at 6. The Times also alleges that Defendants acted unlawfully “[b]y building training datasets containing millions of copies of Times Works” to train their models, and by “storing, processing, and reproducing th[ose] training datasets” as well as “the GPT models trained on Times Works.” Compl. ¶¶ 160-64. Those models include not just GPT-3.5 and GPT-4, but also GPT, GPT-2, and GPT-3. *Id.* ¶¶ 83-92.

2. OpenAI Refuses to Provide Discovery into the Model Training Process (RFP Nos. 1, 3, 6, 8, 14).

OpenAI has refused to provide discovery about key aspects of the model training process. Defendants’ LLMs were developed in stages. Compl. ¶¶ 83-92, 161-62. These stages include: (1) scraping and compiling datasets of works for model training (including Times works), (2) storing and reproducing these training datasets, including scoring, classifying, and cleaning up the works in the training datasets, (3) weighting, mixing, and ordering the works in the training datasets to pass through the model according to a training curriculum, and (4) fine-tuning the models on smaller datasets to work on specific tasks. *Id.* The fine-tuning process also incorporates a technique called Reinforcement Learning from Human Feedback (“RLHF”), in which human reviewers “provide demonstrations of the desired model behavior.” *Aligning Language Models, supra* n.3.

Without precisely defining the limitation it is imposing, OpenAI refuses to produce documents or permit inspection of data for any aspect of training aside from the so-called “pre-training” stage, which at a minimum excludes discovery into fine-tuning and RLHF. Ex. B at 8, 10, 14, 19; Ex. E at 6. This dispute affects the same RFPs addressed above: 1, 3, 6, 8, and 14.

³ *Aligning Language Models to Follow Instructions*, OPENAI (Jan. 27, 2022), <https://openai.com/index/instruction-following/>; *Are GPT-3.5 Series Models Based on GPT-3?*, STACK EXCHANGE, <https://ai.stackexchange.com/questions/39023/are-gpt-3-5-series-models-based-on-gpt-3> (citing an OpenAI blog post that appears to have been taken down).

⁴ *OpenAI API*, OPENAI (June 11, 2020), <https://openai.com/index/openai-api/>.

June 3, 2024

Page 3

All aspects of model training are relevant to the misuse of Times content during these processes, OpenAI's knowledge of that misuse, how Times content (and news content generally) improves model performance, and why Times content shows up later as output. These issues are squarely implicated by the Complaint, in which The Times alleges that Defendants used Times content in various training phases, including the "fine-tuning" phase, and that Defendants prioritized high-quality content for training. Compl. ¶¶ 83-92. Among other things, The Times is entitled to know any preferences or guidelines employed by OpenAI's human reviewers as well as about OpenAI's purported efforts to "monitor for misuse" during fine-tuning, including whether OpenAI previously identified the wrongful conduct at issue in this case. *Aligning Language Models*, *supra* n.3.

OpenAI has offered to search for Times content used for what OpenAI refers to as "post-training" and make available "any such data found as a result of that search." Ex. E at 2. That is insufficient for at least three reasons. First, The Times is not required to accept at face value OpenAI's self-serving assertions about whether Times content was used for any stage of training; The Times is "entitled to discovery to test" those "assertion[s]." *In re Wireless Tel. Servs.*, 2004 WL 764833, at *1 (S.D.N.Y. Apr. 9, 2004). Second, even if Times works were only ingested at certain stages of training, The Times would be entitled to discovery into the entire process, including so its experts can assess how the ingestion of Times works at Point A is relevant to Points B, C, and so on, as well as to determine what Defendants could have done to exclude Times works from the entire process. Third, even if OpenAI makes available Times content it used for post-training, The Times would still need discovery into all of the content used for post-training to understand how OpenAI valued Times content in relation to other content. Finally, OpenAI's argument that discovery into "post-training" would be "unduly burdensome" (Ex. E at 2) is both unsupported and contrary to common sense. OpenAI can easily locate documents that shed light on this aspect of training, and OpenAI has implicitly admitted that it is already reviewing the "post-training" datasets. *Id.*

3. OpenAI Refuses to Provide Discovery About Specific Products (RFP Nos. 5, 12, 13).

OpenAI is likewise excluding from discovery specific products identified in the Complaint, including Bing Chat and Copilot (Compl. ¶¶ 2, 11, 95, 112-17), which OpenAI asserts are "third-party" products. Ex. E at 3. They are not: they are products released by OpenAI's co-defendant in this case (Microsoft), who The Times alleges acted in concert with OpenAI as a joint infringer to develop and commercialize the at-issue models, products, and services. Compl. ¶¶ 93-97, 162, 179. Given these companies' self-described "partnership" for generative AI, OpenAI likely has non-public information about products claimed by Microsoft. Compl. ¶ 94. This dispute affects RFP 5 (seeking documents reflecting policies, procedures or practices concerning the use of intellectual property in AI models) and RFPs 12 and 13 (seeking documents concerning Retrieval Augmented Generation ("RAG"), the process by which Defendants' search applications copy information from the Internet); Compl. ¶¶ 108-23; Ex. B at 11, 17-18. Searching for these plainly relevant documents is not "overly burdensome" (Ex. E at 3), as the parties can negotiate custodians and formulate reasonable search terms based on hit counts.

June 3, 2024
Page 4

Respectfully submitted,

/s/ Ian B. Crosby

Ian B. Crosby
Susman Godfrey L.L.P.

/s/ Steven Lieberman

Steven Lieberman
Rothwell, Figg, Ernst & Manbeck

cc: All Counsel of Record (via ECF)