UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE:	:	
	:	
OPENAI, INC.,	:	25-md-3143 (SHS) (OTW)
COPYRIGHT INFRINGEMENT LITIGATION,	:	
	:	
	:	**ORDER**
This Document Relates To:	:	
All Actions	:	
	:	

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

The Court is in receipt of ECF Nos. 717 and 719 regarding the parties' dispute as to the production of 20 million retained, anonymized consumer ChatGPT output logs (the "Consumer ChatGPT Logs"). (*See* ECF Nos. 717, 719).

At the October 29, 2025, discovery status conference, I directed the parties to submit supplemental briefing articulating (1) the nature of the dispute regarding the production of the 20 million Consumer ChatGPT Logs and (2) whether and how the current dispute relates to News Plaintiffs' prior concerns regarding spoliation. (*See* ECF 715 at 1). The parties filed their supplemental briefs on October 30, 2025. (ECF Nos. 717, 719).

Whether or not the parties had reached agreement to produce the 20 million Consumer ChatGPT Logs in whole—which the parties vehemently dispute—such production here is appropriate. OpenAI has failed to explain how its consumers' privacy rights are not adequately protected by: (1) the existing protective order in this multidistrict litigation or (2) OpenAI's exhaustive de-identification of all of the 20 million Consumer ChatGPT Logs.[1] Further, OpenAI

[1] As News Plaintiffs point out, OpenAI "has spent the last two and a half months processing and deidentifying this 20 million record sample." (ECF 719 at 1 n.1).

has consistently relied on the use of Cochran's Formula by Judge Keulen in *Concord Music Group, Inc. v. Anthropic PBC*, 24-CV-3811 (EKL) (SVK), 2025 WL 1482734 (N.D. Cal. May 23, 2025), in support of its previous proposed methodology for conversation data sampling, (*see e.g.,* ECF 104), but fails to explain why Judge Keulen's subsequent order directing production of the entire 5-million record sample to the plaintiff in that case is not similarly instructive here. *See*, *Concord Music Group, Inc. v. Anthropic PBC*, 24-CV-3811 (EKL) (SVK), 2025 WL 2267950, at *2 (N.D. Cal. Aug. 8, 2025) ("Finally, while the Court is persuaded by concerns relating to Publishers' need for privacy and effectiveness in reviewing the 5-million record sample, the Court finds that Anthropic's proposed compromise largely addresses these concerns.").

Accordingly, News Plaintiffs' motion to compel production of the 20 million de-identified Consumer ChatGPT Logs (ECF 656) is **GRANTED**. OpenAI is directed to produce the 20 million de-identified Consumer ChatGPT Logs to News Plaintiffs by **November 14, 2025**, or within 7 days of completing the de-identification process**.** The parties are directed to continue meeting and conferring regarding any additional measures that the parties can take to effectively protect the privacy interests of OpenAI's consumers.

The Clerk of Court is respectfully directed to close ECF 656.

SO ORDERED.

s/ Ona T. Wang

Dated: November 7, 2025
New York, New York

Ona T. Wang
United States Magistrate Judge