UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE: OPENAI, INC., COPYRIGHT INFRINGEMENT LITIGATION This Document Relates To: THE NEW YORK TIMES COMPANY v. MICROSOFT CORPORATION, et al.	Civil Action No. 25-md-3143

PLAINTIFF'S RESPONSE TO DEFENDANT OPENAI'S OBJECTION TO NON-DISPOSITIVE ORDER PURSUANT TO FEDERAL RULE OF CIVIL PROCEDURE 72(a) (DKT. 638)

TABLE OF CONTENTS

INTRODUCTION 1

BACKGROUND 3

I. The Times Has Complied with its Discovery Obligations 3

A. The Times has already produced significant discovery on its use of Defendants' generative AI products. 3

B. The Times never agreed to produce ChatExplorer logs. 4

II. Magistrate Judge Wang's Well-Reasoned September 19, 2025 Order was Consistent with Governing Law. 6

LEGAL STANDARD 8

ARGUMENT 9

I. OpenAI's Request is Not Relevant to Assessing the Transformativeness of OpenAI's Unauthorized Uses of Times Content. 10

A. ChatExplorer logs are irrelevant to the transformation inquiry. 10

B. *Google v. Oracle* does not support OpenAI's discovery request for ChatExplorer logs. 12

II. OpenAI's Request is Not Relevant to the Fourth Fair Use Factor. 15

A. ChatExplorer logs are irrelevant to the market harm inquiry. 15

B. ChatExplorer logs are irrelevant to the public benefits inquiry. 16

III. ChatExplorer Logs are Not Relevant to OpenAI's Purported Substantial Non-Infringing Uses Defense. 17

IV. OpenAI's Request for ChatExplorer Logs is Disproportionate to the Needs of the Case 19

V. The Court Did Not Clearly Err in Denying OpenAI a 30(b)(6) witness. 22

CONCLUSION 22

TABLE OF AUTHORITIES

Page(s)

Cases

Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith,
598 U.S. 508 (2023).....11

Associated Press v. Meltwater U.S. Holdings, Inc.,
931 F. Supp. 2d 537 (S.D.N.Y. 2013).....12

Authors Guild, Inc. v. HathiTrust,
755 F.3d 87 (2d Cir. 2014).....14

Authors Guild v. Google, Inc.,
804 F.3d 202 (2d Cir. 2015).....14

Coventry Cap. US LLC v. EEA Life Settlements Inc.,
333 F.R.D. 60 (S.D.N.Y. 2019) 8, 9

Gonzales v. National Broad. Co., Inc.,
194 F.3d 29 (2d Cir. 1999).....21

Google LLC v. Oracle America, Inc.,
593 U.S. 1 (2021)..... *passim*

Hachette Book Group., Inc. v. Internet Archive,
115 F. 4th 163 (2d Cir. 2024) 11, 15, 16

In re Keurig Green Mountain Single-Serve Coffee Antitrust Litigation,
341 F.R.D. 474 (S.D.N.Y. 2022) 22

In re OpenAI, Inc., Copyright Infringement Litigation,
No. 23-cv-08292 (SHS), 2025 WL 1652110 (S.D.N.Y. May 30, 2025) 1, 10, 13

Kumaran v. Vision Financial Markets, LLC,
2022 WL 17540669 (S.D.N.Y. Dec. 6, 2022) 9

Madi v. United Parcel Serice General Service Co.,
1999 WL 193403 (S.D.N.Y. Apr. 7, 1999).....8

Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.,
545 U.S. 913 (2005).....18

New York Times Co. v. Microsoft Corp.,
777 F. Supp. 3d 283 (S.D.N.Y. 2025).....18

Nikkal Industries, Ltd. v. Salton, Inc.,
689 F. Supp. 187 (S.D.N.Y. 1988) 9

O'Neill v. Oakgrove Construction, Inc.,
71 N.Y.2d 521 (1988) 21

R.F.M.A.S., Inc. v. So,
748 F. Supp. 2d 244 (S.D.N.Y. 2010).....9

Shand v. Kijakazi,
2023 WL 8379173 (S.D.N.Y. Dec. 4, 2023) 14

United States v. Uccio,
940 F.2d 753 (2d Cir. 1991).....13

Winfield v. City of New York,
2017 WL 5054727 (S.D.N.Y. Nov. 2, 2017).....14

Winfield v. City of New York,
2018 WL 840085 (S.D.N.Y. Feb. 12, 2018).....12

Statutes

17 U.S.C. § 107(1) 10

17 U.S.C. § 107(4) 15

Plaintiff The New York Times Company ("The Times") respectfully submits this opposition to OpenAI's request under Federal Rule of Civil Procedure 72(a) to modify or set aside Magistrate Judge Wang's September 19, 2025 Order. *See* Dkt. 604. As set forth below, Judge Wang's well-reasoned decision should be affirmed.

INTRODUCTION

OpenAI again attempts to frame this lawsuit as being about The Times's use of generative AI, not its own massive copyright infringement. But as this Court has already held, the relevant "use" in this litigation is OpenAI's copying of Times content, *see In re OpenAI, Inc., Copyright Infringement Litig.*, No. 23-cv-08292 (SHS), 2025 WL 1652110, at *4 (S.D.N.Y. May 30, 2025), which includes: (1) OpenAI's copying of The Times's works to create internal datasets; (2) OpenAI's copying of The Times's works to train GPT models; (3) OpenAI's copying of The Times's works to ground results for GPT consumers; and (4) OpenAI's reproduction, abridgment, hallucination, and misattribution of The Times's works in outputs that its GPT products display to those third-party consumers. As Magistrate Judge Wang correctly held, logs from ChatExplorer (a tool used by Times employees that incorporates OpenAI's API) are unrelated to any of those uses.

OpenAI now appeals that ruling, seeking to compel the production of 82,150 ChatExplorer logs from 512 different Times employees by claiming that these logs are "powerful party admissions that OpenAI's LLMs are transformative." Dkt. 639 at 1. They are no such thing. To the extent The Times's admission of its *use* of ChatExplorer is relevant (it is not), OpenAI already has such admissions. *See* Ex. A (The New York Times Company's Responses and Objections to OpenAI's Fourth Set of Requests for Admission). Indeed, as The Times has already explained in its prior Rule 72(a) opposition briefs, it has produced documents regarding its use of Defendants' LLMs in reporting or presenting content, its trainings and policies regarding Defendants' LLMs

(including ChatExplorer), documents relating to the AI Initiatives program, and Audience Insight Group analyses—which directly assess the impact of Generative AI on The Times. And OpenAI has questioned Times witnesses extensively about their use of both ChatExplorer and the consumer-facing ChatGPT. *See, e.g.*, Ex. B (Joy Robins Dep. Tr.) at 232:5–234:20; Ex. C (Bailey Evans Dep. Tr.) at 184:9–185:22.

The contents of the Times's ChatExplorer logs will not add anything further. Specifically, ChatExplorer logs only show how Times employees interacted with OpenAI's API—and Times employees by definition could not infringe their own copyrights, nor are they especially probative of how users interact with OpenAI's consumer-facing products in the real world. In contrast, OpenAI already has billions of logs in its possession, which will demonstrate the ways in which its ***800 million weekly active users*** prompt its products and whether such outputs are infringing or otherwise substitutive. Yet, while opposing The Times's motion to compel those very logs—which are the most direct evidence of infringing GPT product outputs—OpenAI now seeks 82,150 ChatExplorer logs from The Times. *See* Dkt. 717 (OpenAI's Letter Opposition to News Plaintiffs' Pending Motion to Compel OpenAI's Output Log Data). That is absolutely improper and not a reasonable basis for overturning Magistrate Judge Wang's well-reasoned decision.

OpenAI acknowledges "that objections to a magistrate judge's discovery order are reserved for rare circumstances." Dkt. 639 at 1. Nonetheless, for the third time, OpenAI files yet another Rule 72(a) objection challenging a discovery order issued by Magistrate Judge Wang. Not only does this waste the Court's resources, but it also disregards the extensive efforts Magistrate Judge Wang has put into resolving these discovery disputes. Because OpenAI has not met its exacting burden of demonstrating that Judge Wang's decision is clearly erroneous or contrary to law, its Objection should be denied.

BACKGROUND

I. The Times Has Complied with its Discovery Obligations.

A. The Times has already produced significant discovery on its use of Defendants' generative AI products.

OpenAI's effort to overturn Judge Wang's order continues its trend of demanding discovery that is irrelevant, unduly burdensome, and disproportionate to the needs of the case. Specifically, Defendants sought discovery of 82,150 ChatExplorer logs from 512 different Times employees. Soon after making this request, OpenAI then sought to compel logs from ***fifteen other tools*** that The Times uses. Defendants' primary basis for seeking these logs was that it needed evidence to show "that the accused products are used for transformative purposes." Dkt. 396 at 3.

Because that argument is untethered to the fair use analysis—which focuses on the "purpose" for which a defendant copied a plaintiff's work—The Times has always disagreed that discovery into its use of Defendants' generative AI tools is relevant to this case. Nevertheless, in an effort to move the case along, The Times produced extensive documentation on its uses of Defendants' products. Specifically, The Times agreed to produce the following categories of documents:

1. Documents regarding The Times's use of Defendants' generative AI tools in reporting or presenting content;[1]
2. Documents regarding The Times's trainings and policies regarding the use of Defendants' generative AI products;
3. Documents relating to the "AI Initiatives" program to evaluate and develop potential uses of AI at The Times;

[1] To be clear, The Times does not use generative AI of any kind to draft its articles. *See* https://www.nytimes.com/2024/10/07/reader-center/how-new-york-times-uses-ai-journalism.html.

4. Documents regarding how generative AI can benefit The Times's audience; and
5. Audience Insight Group (AIG) analyses, which directly assess the impact of generative AI on The Times.

These custodial documents are the best evidence of The Times's use of Defendants' generative AI tools—if any—in reporting or presenting content, as well as The Times's guidance regarding the permitted and unpermitted uses of Defendants' products. As just one example, The Times has produced multiple AI Initiatives presentations, which discuss in detail The Times's ongoing research into Generative AI products and their potential use to enhance The Times's offerings. Similarly, The Times has provided swathes of guidance regarding ChatExplorer, which describe the tool in detail, its permitted and unpermitted uses, and different GPT models that it employs. The Times has even agreed to produce AIG reports through October 2025, which include its own internal assessment of the impact of Generative AI on The Times.

The Times has also provided multiple admissions to Defendants' Requests for Admissions regarding its use of ChatExplorer. *See* Ex. A. And, OpenAI has questioned Times witnesses extensively about their use of both ChatExplorer and the consumer-facing ChatGPT. *See, e.g.*, Ex. B (Joy Robins Dep. Tr.) at 232:5–234:20; Ex. C (Bailey Evans Dep. Tr.) at 184:9–185:22. Finally, The Times has agreed to provide 30(b)(6) testimony on the uses of Defendants' tools for reporting and the presentation of content. Given this extensive production, producing ChatExplorer logs for 512 Times employees is neither relevant nor proportional to the needs of this litigation.

B. The Times never agreed to produce ChatExplorer logs.

ChatExplorer, which is The Times's internal interface of the OpenAI API, became generally available to Times employees on December 14, 2023. *See* Dkt. 417-1 ¶ 3. It is an optional

tool for experimentation by Times employees that they were not required to use. The Times eventually stopped adding users to ChatExplorer.

The Times promptly disputed the relevance of ChatExplorer logs when OpenAI first sought them in February and March 2024, asserting that ChatExplorer logs have no connection to OpenAI's defenses and would require substantial review that is unwarranted given their lack of relevance. Notwithstanding its relevance objection, The Times's in-house counsel directed a hold be placed on ChatExplorer logs. *Id.* ¶ 2.

The Times employee tasked with implementing this hold mistakenly believed those logs were already under a one-year preservation policy. *Id.* ¶ 3. However, at the time, ChatExplorer data was inadvertently being stored in a cloud-based platform that had a 90-day retention period. *Id.* ¶ 4. When this was discovered on June 20, 2024, the retention policy on the cloud partition was immediately changed to 180 days to prevent further data loss and the logs were ultimately stored with a 10-year retention period. *Id.* ¶¶ 4–5. As of June 2025, The Times had approximately 82,150 logged requests saved.

Despite its willingness to preserve this data—and its proactive steps to do so—The Times never agreed to produce it. Instead, in an effort to move litigation forward, The Times offered to provide hit counts for a targeted set of search terms run over a set of specific custodians so that The Times could consider OpenAI's production request; thus, it asked OpenAI to propose terms for a subset of custodians. OpenAI instead demanded all logs for 58 different employees—many of whom were not custodians in this litigation. The Times declined OpenAI's overbroad request and provided log counts for its custodians. It made no representations about production, given OpenAI's refusal to even provide terms for The Times to run. Instead of negotiating with The Times, OpenAI doubled down on its request and demanded that The Times produce ***all*** 82,150

ChatExplorer logs from 512 different employees. When The Times refused, OpenAI and Microsoft filed motions to compel all ChatExplorer logs. *See* Dkt. 377; 396.

II. Magistrate Judge Wang's Well-Reasoned September 19, 2025 Order was Consistent with Governing Law.

On July 29, 2025, OpenAI moved to compel production of all 82,150 ChatExplorer logs and requested a four-hour 30(b)(6) deposition regarding The Times's data preservation. *See* Dkt. 377. Microsoft also moved to compel, arguing that ChatExplorer logs were relevant to Defendants' substantial non-infringing use defense. *See* Dkt. 396. Judge Wang heard argument on Defendants' motions at the August 12, 2025 discovery hearing. Following the hearing, the Court ordered supplemental briefing on the relevance of ChatExplorer logs, as well as the proportionality and burden associated with their production. The parties filed supplemental briefing on August 22, 2025. *See* Dkt. 472; Dkt. 475; Dkt. 477.

After carefully considering the parties' submissions and oral arguments, Judge Wang denied Defendants' motions to compel, holding that Defendants failed to demonstrate the "relevance or proportionality of the discovery sought." Dkt. 604 at 1. In reaching this conclusion, the Court first noted that The Times "produced 'significant discovery' on [The Times's] own use of Defendants' LLMs, including documents regarding: (1) [The Times's] use of Defendants' LLMs in reporting or presenting content; (2) [The Times's] trainings and policies re Defendants' LLMs (including guidance relating to ChatExplorer); (3) [The Times's] AI Initiatives program; and (4) Audience Insight Group reports that assess the impact of GenAI on [The Times]." *Id.* at 2.

With that in mind, the Court considered whether ChatExplorer logs were relevant to the first fair use factor and correctly noted that "the relevant inquiry under the first use factor concerns a defendant's use of a plaintiff's copyrighted material, not a downstream use of defendant's allegedly infringing material by a copyright-holder plaintiff." *Id.* at 5–6 (emphasis in original).

The Court explained that the "relevant 'uses' at issue in this case are: (1) Defendants' alleged copying of [The Times's] and other plaintiffs' copyrighted works to (a) create an internal dataset and (b) train GPT models; and (2) the Defendants' alleged copying of Plaintiffs' copyrighted material in the outputs of Defendants' LLMs." *Id.* at 6. Given that Defendants did "not explain how the content of [The Times's] ChatExplorer logs would have any bearing on how the uses of Defendants' LLMs might or might not be transformative or commercial," the Court concluded that Defendants' requested discovery was irrelevant to the first fair use factor. *Id.*

Next, the Court stated that "Defendants d[id] not explain how the content of the ChatExplorer [l]ogs would support a conclusion of either a public benefit to nonparties or any conclusion of market harm." *Id.* at 7. With respect to market harm, the Court correctly noted that "[t]he fourth fair use factor requires an analysis of the Defendants' use of [The Times's] copyrighted material—not [The Times's] use of Defendants' LLMs—and whether Defendants' use impacts the market for [The Times's] copyrighted works." *Id.* at 8. The Court determined that The Times's "use of ChatExplorer is not relevant to whether other users of Defendants' LLMs found in those LLMs a 'competing substitute' for [The Times's] copyrighted works." *Id.* (emphasis in original). That is because The Times "cannot be a 'competing substitute' for itself or deprive itself of revenue." *Id.*

With respect to the public benefits inquiry, the Court noted that it "requires a court to 'balance the benefit the public will derive if the use is permitted against the personal gain the copyright owner will receive if the use is denied.'" *Id.* (emphasis in original) (quoting *Hachette Book Grp., Inc. v. Internet Archive*, 115 F. 4th 163, 195 (2d Cir. 2024)). The Court concluded that "any benefit to [The Times] from its use of ChatExplorer, even it could be discovered or inferred from the content of the ChatExplorer logs, would be probative only of a potential benefit to [The

Times] and would not be probative of any public benefit." *Id.* at 7 (emphasis in original). Accordingly, the Court held that "ChatExplorer logs are not probative of [The Times's] claims of market harm or the alleged public benefit of Defendants' LLMs, and thus are not relevant to factor four of the fair use analysis under Rule 26." *Id.* at 9 (footnote omitted).

The Court also found that ChatExplorer logs were not relevant to Defendants' purported substantial non-infringing uses defense because The Times "is the copyright holder and cannot infringe its own copyrights, so every use by [The Times] is, by definition, non-infringing." *Id.* at 11 (emphasis in original). Thus, the Court concluded that The Times's "use of Defendants' LLMs [ ] cannot be probative of the existence of substantial non-infringing uses, nor the commercial significance of such uses." *Id.*

Given that "Defendants [ ] failed to sustain their burden of showing the relevance of the ChatExplorer logs, the Court [noted that it] need not reach proportionality." *Id.* However, the Court stated that even if the logs were relevant, "ordering their production would still not be proportional to the needs of this case," as The Times "would need to individually review each of the 80,000+ log entries for attorney client privilege and reporter's privilege." *Id.* at 12.

Finally, "[b]ecause the ChatExplorer logs are not relevant," the Court denied OpenAI's request for a special 30(b)(6) deposition.

LEGAL STANDARD

Under Federal Rule of Civil Procedure 72(a), "a magistrate judge's discovery ruling is a non-dispositive order which can only be overturned if it is 'clearly erroneous or contrary to law.'" *Madi v. United Parcel Serv. Gen. Serv. Co.*, 1999 WL 193403, at *1 (S.D.N.Y. Apr. 7, 1999) (quoting Fed. R. Civ. P. 72(a)). "A finding is clearly erroneous if the reviewing court is left with the definite and firm conviction that a mistake has been committed." *Coventry Cap. US LLC v.*

EEA Life Settlements Inc., 333 F.R.D. 60, 64 (S.D.N.Y. 2019) (quoting *R.F.M.A.S., Inc. v. So*, 748 F. Supp. 2d 244, 248 (S.D.N.Y. 2010)). "Similarly, a finding is contrary to law when it fails to apply or misapplies relevant statutes, case law, or rules of procedure." *Id.* (quoting *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248).

"Therefore, pursuant to these standards, a magistrate judge's determinations on discovery matters are entitled to substantial deference." *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248; *see Nikkal Indus., Ltd. v. Salton, Inc.*, 689 F. Supp. 187, 189 (S.D.N.Y. 1988) ("Consistently, it has been held that a magistrate's report resolving a discovery discourse between litigants should be afforded substantial deference and be overturned only if found to be an abuse of discretion."). "The magistrate judge's findings should not be rejected merely because the court would have decided the matter differently." *Kumaran v. Vision Fin. Markets, LLC*, 2022 WL 17540669, at *6 (S.D.N.Y. Dec. 6, 2022). "The party seeking to overturn a magistrate judge's decision thus carries a heavy burden." *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248 (quoting *U2 Home Ent., Inc. v. Hong Wei Int'l Trading Inc.*, 2007 WL 2327068, at *1 (S.D.N.Y. Aug. 13, 2007)).

ARGUMENT

OpenAI has failed to satisfy its heavy burden of demonstrating that Magistrate Judge Wang's September 19, 2025, Order was clearly erroneous or contrary to law. *First*, the Court correctly held that ChatExplorer logs were not relevant to assessing whether OpenAI's unauthorized use of The Times's works was transformative. *Second*, the Court correctly ruled that ChatExplorer logs were not relevant to the market harm or public benefit inquiries—that is, the fourth statutory fair use factor. *Third*, the Court properly determined that ChatExplorer logs were not relevant to Defendants' putative defense of substantial non-infringing use—which is not, in fact, at issue in this case. *Fourth*, the Court appropriately determined that reviewing and producing

82,150 ChatExplorer logs was disproportionate to the needs of the case. Magistrate Judge Wang's Order should be affirmed.

I. OpenAI's Request is Not Relevant to Assessing the Transformativeness of OpenAI's Unauthorized Uses of Times Content.

First, OpenAI contends that the Court clearly erred in ruling that "the only 'use' relevant to evaluating whether OpenAI's use is 'transformative' is ***OpenAI's*** use of The Times's works." Dkt. 639 at 9 (emphasis in original). OpenAI asserts that "ChatExplorer logs are uniquely relevant evidence of the purposes and capabilities of OpenAI's LLMs," as "[t]hey are party admissions from The Times . . . [that] will show exactly how those LLMs can be and are used for new purposes." *Id.* at 10. This argument fails for several reasons.

A. ChatExplorer logs are irrelevant to the transformation inquiry.

The issue of transformativeness is part of the first fair use factor, which assesses "the purpose and character of the use, including whether such use is of a commercial nature." 17 U.S.C. § 107(1). As this Court has previously held, "the first fair use factor is concerned with '*the copier's use* of an original work,' and 'the extent to which *the secondary use* is transformative.'" *In re OpenAI, Inc., Copyright Infringement Litig.*, No. 23-cv-08292 (SHS), 2025 WL 1652110, at *4 (emphases in original) (first quoting *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528 (2023); and then quoting *Hachette Book Grp., Inc. v. Internet Archive*, 115 F. 4th 163, 179 (2d Cir. 2024)). In making this assessment, courts consider "whether *the copier's* use 'adds something new, with a further purpose or different character, altering' the copyrighted work 'with new expression, meaning or message.'" *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 29 (2021) (emphasis added) (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994)).

"This litigation is focused on [D]efendants' allegedly infringing use of [P]laintiffs' works, and [D]efendants' fair use defense is likewise grounded in that use." *In re OpenAI, Inc., Copyright*

Infringement Litig., No. 23-cv-08292 (SHS), 2025 WL 1652110, at *4. Accordingly, the relevant analysis focuses on what OpenAI wanted to achieve through its copying, including whether it sought out Times works because they could help consumer tools like ChatGPT deliver news with better accuracy and expression—the same purpose for which people read The Times's journalism. *See Andy Warhol Found.*, 598 U.S. at 526 (finding that Warhol's use of Goldsmith's photograph for a magazine cover was not "transformative" because "[a]s portraits of Prince used to depict Prince in magazine stories about Prince, the original photograph and [Warhol's] copying use of it share substantially the same purpose"). The Times's use of OpenAI's products is entirely unrelated to that inquiry. *See id.* at 509 (noting that the first fair use factor "considers the reasons for, and nature of, the copier's use of an original work"); *Hachette Book Group Inc.*, 115 F.4th at 179 (noting that in assessing the first fair use factor, courts consider "the extent to which the secondary use is transformative").

Instead, OpenAI offers two theories to compel the production of the logs that are both disconnected from the fair use inquiry. First, OpenAI asserts that ChatExplorer logs "are powerful party admissions that OpenAI's LLMs are transformative." Dkt. 639 at 1. However, OpenAI fails to demonstrate how *the contents of ChatExplorer logs* constitute "party admissions" or are relevant to the transformative inquiry. To the extent The Times's admission of its *use* of ChatExplorer is relevant, OpenAI already has such admissions. *See* Ex. A at 4, 6 (The New York Times Company's Responses and Objections to OpenAI's Fourth Set of Requests for Admission) (admitting "that ChatExplorer was a user interface [The Times] used internally"); (admitting "that ChatExplorer used OpenAI's GPT 4 Large Language Model"). Moreover, The Times has already produced documents regarding its use of OpenAI's products in reporting or presenting content and its trainings and policies regarding ChatExplorer. The contents of the Times's ChatExplorer logs will

not provide anything further. Specifically, OpenAI already has billions of logs in its possession—including, presumably, identical logs of the very same interactions with its own products that it is seeking from The Times—which will demonstrate the ways in which its ***800 million weekly active users*** prompt its products.[2] There is nothing further that ChatExplorer logs—which only show how Times employees may have prompted OpenAI's API—will add to the transformativeness inquiry.[3]

Second, OpenAI asserts that it is entitled to ChatExplorer logs because The Times itself has sought discovery into OpenAI's output log data. *See* Dkt. 639 at 14. This is a red herring. OpenAI's output log data is relevant because it serves as direct evidence of OpenAI's copyright infringement. ChatExplorer logs do not.

Accordingly, given that OpenAI "bears the initial burden of demonstrating relevance and proportionality," its failure to articulate an adequate theory of relevance is fatal to its request for ChatExplorer logs. *Winfield v. City of New York*, 2018 WL 840085, at *3 (S.D.N.Y. Feb. 12, 2018).

B. *Google v. Oracle* does not support OpenAI's discovery request for ChatExplorer logs.

As it did in its last two unsuccessful Rule 72(a) objections, OpenAI once again misreads *Google v. Oracle*. OpenAI claims that in finding Google's secondary use of Oracle's API to create the Android platform transformative, the Supreme Court "relied on the fact that Android was a highly creative and innovative tool for a smartphone environment that could be used by engineers and developers." Dkt. 639 at 10. Based on this generalized observation, OpenAI concludes that

[2] *See* https://openai.com/index/how-people-are-using-chatgpt/

[3] The Times also continues to dispute the transformativeness of OpenAI's products. The transformative "inquiry asks not merely whether the new work is transformative, but also the extent to which it transforms the copyrighted work." *Associated Press v. Meltwater U.S. Holdings, Inc.*, 931 F. Supp. 2d 537, 551 (S.D.N.Y. 2013). There is nothing transformative about the ***infringing*** use at issue here, which merely ***consumes*** the works for the information and expression they contain to enable the models to output similar works that dilute the market for the originals.

the Supreme Court endorsed the broad view that a plaintiff's use of a defendant's product is always relevant to the fair use inquiry. *See id.*

The Court has already rejected this overly broad reading of *Google*. *In re OpenAI, Inc., Copyright Infringement Litig.*, No. 23-cv-08292 (SHS), 2025 WL 1652110, at *3 ("*Oracle* does not stand for the proposition that evidence of the plaintiff's activities is *always* relevant."); *see also Google*, 593 U.S. at 36 ("We do not say that these questions are always relevant to the application of fair use."). It should do so again. *See United States v. Uccio*, 940 F.2d 753, 758 (2d Cir. 1991) ("[W]hen a court has ruled on an issue, that decision should generally be adhered to by that court in subsequent stages in the same case.").

However, to the extent the Court is willing to again entertain OpenAI's argument, it is unavailing for the very reasons the Court has previously identified. In *Google*, the Supreme Court considered whether Google's copying of a copyrighted computer program developed by Sun—which Oracle acquired in 2010—constituted a fair use. In its discussion of the first fair use factor, the Supreme Court noted that the record "demonstrate[d] the numerous ways in which" Google's use was transformative. Specifically, the Court stated that the jury heard about the specific features and functionalities of the Android platform. Crucially, however, the Court **did not** list Java's use of the Android platform as relevant evidence in assessing the transformative nature of the Android platform. If such evidence were truly as "critical" as OpenAI asserts, it should have been cited by the *Google* Court in supporting its conclusion. However, at no point in its decision, did the Supreme Court mention that such evidence existed in the record or that it was relevant to the Supreme Court's decision to find Google's secondary use transformative. Accordingly, *Google*

does not stand for the proposition that The Times's use of OpenAI's LLMs was relevant to the first fair use factor.[4]

The Supreme Court also noted that "Sun executives thought that widespread use of the Java programming language . . . would benefit the company." *Google*, 593 U.S. at 31. But here, The Times alleged that widespread use of Defendants' products pose a significant threat to The Times' business and has produced discovery relevant to that contention. *See, e.g.*, NYT_01018244 (AIG September 2025 GenAI Research Summary) ("By now, most of our target audiences, and especially our readers and subscribers, are using GenAI. At the low end, 74% of US news consumers say they use it to get news and 93% of US sports media consumers say they use it for sports at the high end. This gives us strong conviction that GenAI is taking a share of engagement among our current and potential readers and subscribers across all of our products." (emphasis in original)). Accordingly, OpenAI's reliance on *Google* is inapt.

Finally, as The Times has previously noted, OpenAI's reliance on the *Google* Court's observations fails for an additional reason: the Supreme Court was not ruling on whether any evidence was discoverable or admissible. *Google*, 593 U.S. at 31–32. Thus, any language bearing on those inquiries is merely dicta. *See Shand v. Kijakazi*, 2023 WL 8379173, at *2 n.1 (S.D.N.Y. Dec. 4, 2023) ("It is well understood that 'dicta' are the '[t]he things the court said, not what it held; and only what a court holds is binding . . . in subsequent cases.'" (alteration and omission in original) (internal quotation marks omitted)). Comments made in *dicta* cannot establish that a court's ruling was clearly erroneous or contrary to the law. *See Winfield v. City of New York*, 2017

[4] For the same reasons, OpenAI's reliance on *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015), and *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014) is unavailing. Specifically, neither opinion stands for the proposition that the downstream use of a defendant's technology by a copyright-holder plaintiff is relevant to the transformative inquiry, nor did the Second Circuit mention the plaintiff's use of defendant's technology as relevant evidence when assessing the first fair use factor.

WL 5054727, at *3 (S.D.N.Y. Nov. 2, 2017) (concluding that the magistrate judge's decision "was not contrary to law" where there was an "absence of any contravening ***binding*** authority" (emphasis added)).

The Court did not clearly err in ruling that ChatExplorer logs were irrelevant to the first fair use factor.

II. OpenAI's Request is Not Relevant to the Fourth Fair Use Factor.

Second, OpenAI contends that the Court clearly erred in ruling that ChatExplorer logs were not "relevant to fair use factor four, which considers the harm to the market for the copyrighted work, as well as the 'public benefits' of the alleged use." Dkt. 639 at 15. However, this argument fails for the very same reasons discussed above.

A. ChatExplorer logs are irrelevant to the market harm inquiry.

The fourth fair use factor assesses "the effect of the use upon the potential market for or value of the copyrighted work." 17 U.S.C. § 107(4). In evaluating this factor, courts analyze "whether there was harm to the actual or potential markets for the copyrighted work." *Google*, 593 U.S. at 26. More specifically, "[t]his factor 'focuses on whether the copy brings to the marketplace a competing substitute for the original, or its derivative, so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may to opt to acquire the copy [rather than] the original.'" *Hachette Book Grp., Inc.*, 115 F.4th at 189 (second alteration in original) (quoting *Authors Guild v. Google, Inc.*, 804 F.3d 202, 223 (2d Cir. 2015)). "[T]he ultimate burden of proving that the secondary use does not compete in the relevant market is appropriately borne by the party asserting the [fair use] defense: the secondary user." *Id.* at 194.

Here, "the use" and "the copy" at issue are OpenAI's unauthorized use and copying of The Times's asserted works, and "the potential market" is the market for the same Times works. In other words, the market harm inquiry assesses the extent to which users of OpenAI's LLMs found

in those LLMs a "competing substitute" for The Times's copyrighted works. ChatExplorer logs are entirely irrelevant to that inquiry. Only Times employees used ChatExplorer, and Times employees' use of ChatExplorer cannot conceivably have an effect on the market for The Times's own works because The Times cannot be a "competing substitute" for itself or deprive itself of revenue. The only logs that are relevant to this inquiry are third parties' logs, which can show how Defendants' LLMs served as a "competing substitute" for The Times's copyrighted works—and those are entirely in the custody, possession, or control of OpenAI.

OpenAI does not meaningfully engage with the market harm inquiry. Instead, OpenAI asserts that "the ChatExplorer logs are likely to provide concrete examples showing that, contrary to The Times' allegations, The Times uses OpenAI's technologies to generate revenue and improve its journalism." Dkt. 639 at 15–16. However, that fundamentally misunderstands the fourth fair use factor. Whether The Times has experimented with OpenAI's LLMs has no bearing on whether OpenAI's LLMs "bring[] to the marketplace a competing substitute for [The Times's copyrighted works]." *Hachette Book Grp., Inc.*, 115 F.4th at 189. Accordingly, OpenAI has failed to demonstrate the relevance of ChatExplorer logs to the market harm inquiry.

B. ChatExplorer logs are irrelevant to the public benefits inquiry.

In evaluating the fourth fair use factor, courts also assess "the public benefits ***the copying*** will likely produce." *Google*, 593 U.S. at 35 (emphasis added). Courts typically balance "the benefit the public will derive if the use is permitted against the personal gain the copyright owner will receive if the use is denied." *Hachette Book Grp.*, 115 F. 4th at 195.

As with the market harm inquiry, the relevant use at issue here is OpenAI's unauthorized use and copy of The Times's copyrighted works. And the public benefits inquiry assesses to what extent ***the public*** benefits from OpenAI's unauthorized use of The Times's works. OpenAI has access to ***billions*** of logs, which it will likely rely on to argue that the public uses and benefits

from its products. Yet, while restricting The Times's access to such logs, OpenAI insists on the production of 82,150 ChatExplorer logs. How The Times—*i.e.,* one of the 800 million users of OpenAI's products—uses OpenAI's products is irrelevant to the broader public benefits inquiry. Indeed, OpenAI does not even attempt to show how these 82,150 logs (as compared to the ***billions*** of logs in OpenAI's possession) are uniquely situated to show the benefits to the broader public. Thus, OpenAI's demand that The Times review and produce its ChatExplorer logs—a tiny amount compared to the vast number of logs to which OpenAI has access—is unreasonable and disproportionate to the needs of this case.

Accordingly, the Court did not clearly err in holding that The Times's ChatExplorer logs are irrelevant to the fourth fair use factor.[5]

III. ChatExplorer Logs are Not Relevant to OpenAI's Purported Substantial Non-Infringing Uses Defense.

Third, OpenAI contends that "[t]he Order should be set aside for the additional reason that it erroneously concluded that the ChatExplorer logs were irrelevant to showing substantial non-infringing uses." Dkt. 639 at 20. Specifically, OpenAI claims that "ChatExplorer logs are relevant to this substantial non-infringing use defense because . . . they provide concrete examples of how OpenAI's technology is used for purposes other than infringing The Times's copyrights." *Id.* at 21. This argument fails for two reasons.

[5] OpenAI also claims that since Magistrate Judge Wang's September 19, 2025 Order issued, "The Times has relied on it to withhold discovery into additional relevant uses" of Defendants' products. Dkt. 639 at 19. As an initial matter, The Times notes that it never conceded that its use of Defendants' products was relevant to the fair use inquiry. Indeed, Judge Wang's Order confirms that it is not. *See* Dkt. 604 at 5 ("As this Court has explained before, the relevant inquiry under the first fair use factor concerns a defendant's use of a plaintiff's copyrighted material, not a downstream use of defendant's allegedly infringing material by a copyright-holder plaintiff."); *id.* at 8 ("The fourth fair use factor requires an analysis of the Defendants' use of NYT's copyrighted material—not NYT's use of Defendant's LLMs—and whether Defendants' use impacts the market for NYT's copyrighted works."). Instead, The Times only agreed to produce such documents to move the litigation forward. In any event, OpenAI only points to The Times's refusal to produce logs for fifteen additional products as evidence of its "withhold[ing] discovery." Dkt. 639 at 20. Those logs are irrelevant for the very reasons ChatExplorer logs are irrelevant.

As an initial matter, the substantial non-infringing uses defense is not at issue in this case. The fact that a product may be used in ways that are not infringing has no relevance to a claim against the uses of the product that are infringing. Rather, the defense of substantial non-infringing use is only relevant when a rights holder asserts a claim of ***induced*** infringement by mere distribution of a product. *See Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 934–35 (2005) ("We agree with MGM that the Court of Appeals misapplied *Sony*, which it read as limiting secondary liability quite beyond the circumstances to which the case applied. *Sony* barred secondary liability based on presuming or imputing intent to cause infringement solely from the design or distribution of a product capable of substantial lawful use, which the distributor knows is in fact used for infringement."). As this Court has recognized, The Times has not asserted such a claim in this case. *See New York Times Co. v. Microsoft Corp.*, 777 F. Supp. 3d 283, 308–09 (S.D.N.Y. 2025) ("[W]hile *Sony* and *Grokster* analyzed claims of contributory copyright infringement by inducement, they did not discuss claims of contributory copyright infringement by material contribution, which plaintiffs allege here. Indeed, 'the fact that a product is capable of substantial lawful use does not mean the producer can never be held contributorily liable.'" (quoting *BMG Rts. Mgmt. (US) LLC v. Cox Commc'ns, Inc.*, 881 F.3d 293, 306 (4th Cir. 2018))).

However, were this defense at issue in this case, ChatExplorer logs would not be relevant because The Times is the copyright holder and cannot infringe its own copyrights. Thus, every use by The Times is—by definition—non-infringing of its own works and, thus, not probative of the existence of substantial non-infringing uses in this matter. Instead, what is relevant is how ***other consumers***—*i.e.*, those who can use OpenAI's products in infringing ways—are using Defendants' LLMs and whether those third-party uses are for substantially non-infringing purposes.

Accordingly, the Court did not clearly err in concluding that ChatExplorer logs were irrelevant to Defendants' substantial non-infringing uses defense.

IV. OpenAI's Request for ChatExplorer Logs is Disproportionate to the Needs of the Case.

Fourth, OpenAI contends that "[t]he Order clearly erred by holding that production of the ChatExplorer logs was not proportionate to the needs of the case." Dkt. 639 at 21. Specifically, OpenAI claims that "The Times made ***no*** showing of undue burden." *Id.* at 22 (emphasis in original). OpenAI is incorrect.

For the reasons already discussed, ChatExplorer logs are not relevant. However, assuming for the sake of argument the Court finds the logs relevant, their production goes against the tenets of Federal Rule of Civil Procedure 26(b)(1). As The Times explained in its briefing to Magistrate Judge Wang, the burden of reviewing 82,150 unique documents outweighs their likely benefit. *See* Dkt. 475 at 2–5. Contrary to OpenAI's assertion, The Times's review process for these documents would be far more involved than scanning a spreadsheet. Instead, exporting the ChatExplorer data to a review platform generates a separate document for each call and response pair, with header information indicating the user. *See* Dkt. 376-1 at 1. The document may be less than a page, or it may span several pages. With regard to the format, these documents are not unlike emails. And, reviewing an additional 82,150 emails would be an unreasonable ask in light of these documents' minimal relevance.

However, review of these logs is significantly more burdensome than custodial email review. While traditional document review is limited to mutually agreed-upon custodians and a tailored set of terms, The Times would be required to review the logs from 512 distinct users, most of whom are not custodians in this case, and that includes members of the newsroom. *See id.* at 2–3. Specifically, these non-custodians are associated with over 81,000 of the 82,150 ChatExplorer

logs. Thus, for a vast majority of these ChatExplorer logs (98.5%), a reviewer will need to go through at least the following steps for each and every unique log to properly understand the context:

1. Determine who the relevant individual is;
2. Determine what that individual's role at The Times entails;
3. Determine what their prompt concerns;
4. Determine whether the prompt includes any sensitive or confidential information; and
5. Determine whether the prompt implicates any privilege (i.e., attorney-client and/or reporters' privilege.)

Much of this information is not readily available to The Times's attorneys, and it would take considerable time to go out and acquire. Most critically, unlike emails, which generally appear in a chain that provides some context, the context of a prompt can often be unclear from the face of the document—meaning reviewers may need to individually meet with users to clarify the answers to items (3), (4), and (5) above. Performing this exercise hundreds, let alone tens of thousands of times, is vastly more burdensome than is justified by any minimal relevance of these documents.

And, the need to review these logs carefully is critical. The Times's ChatExplorer guidance explicitly permits users to input sensitive Times content, such as The Times's "source code." Clearly, The Times's source code is highly sensitive information and references to it would need to be, at minimum, flagged for special treatment. Then, of course, there are privilege concerns. The Times would need to review the logs for inputs that are protected by the attorney-client or reporters' privilege—both of which require understanding the full context of any given input.

Regarding the reporters' privilege in particular, the Second Circuit "has long recognized the existence of a qualified privilege for journalistic information" under federal common law. *Gonzales v. Nat'l Broad. Co., Inc.*, 194 F.3d 29, 32 (2d Cir. 1999). "The qualified reporter's privilege, accorded under the United States Constitution, the New York State Constitution, and the New York Shield law, protects ***non-confidential*** as well as confidential information" "acquired in the course of gathering news." *Application of Behar*, 779 F. Supp. 273, 274–275 (S.D.N.Y. 1991). New York courts have protected news organizations from the burdens of discovery in order to protect the "ability of the press freely to collect and edit news, unhampered by repeated demands for its resource material," noting that the "practical burdens on time and resources, as well as the consequent diversion of journalistic effort and disruption of newsgathering activity, would be particularly inimical to the vigor of a free press." *O'Neill v. Oakgrove Constr., Inc.*, 71 N.Y.2d 521, 526–27 (1988).

The Times would thus need to review newsroom employees' logs to ensure information protected by the reporters' privilege is not being produced. Consistent with The Times's Generative AI guidance, newsroom employees were permitted to prompt ChatExplorer "to help with brainstorming and editing, summarize Times-owned materials and non-proprietary information." In so doing, newsroom employees may have inputted "information acquired in the course of gathering news." *Application of Behar*, 779 F. Supp. at 274. Given that 1,285 individuals completed the ***newsroom*** Generative AI training and thus could have accessed ChatExplorer, there is reason to believe this will be a considerable issue. And while The Times opposes any production of these logs, under no circumstances have Defendants justified the production of logs from individual users associated with The Times's newsroom.

Simply put, producing 82,150 logs means reviewing that many additional unique documents—and, unlike with custodial emails, potentially performing repeated individual investigations into the context of tens of thousands of prompts. In short, "the burden or expense of the proposed discovery outweighs its likely benefit." Fed. R. Civ. P. 26(b)(1).

Accordingly, the Court did not clearly err in concluding that OpenAI's request for ChatExplorer logs was disproportionate to the needs of the case.

V. The Court Did Not Clearly Err in Denying OpenAI a 30(b)(6) witness.

Finally, OpenAI claims that "the Order's denial of OpenAI's request for a limited deposition regarding The Times's destruction of the ChatExplorer logs was also clearly erroneous." Dkt. 639 at 23. However, because ChatExplorer logs are not relevant, the inadvertent loss of these logs does not implicate Federal Rule of Civil Procedure 37(e). *See In re Keurig Green Mountain Single-Serve Coffee Antitrust Litigation*, 341 F.R.D. 474, 497 (S.D.N.Y. 2022) (holding that spoliation claims require the moving party to "show that the lost information was not only probative, but that it would affirmatively support the movant's claim").[6]

Accordingly, the Court did not clearly err in denying OpenAI a 30(b)(6) witness on the inadvertent deletion of some ChatExplorer logs.

CONCLUSION

For the foregoing reasons, The Times respectfully requests the Court affirm Magistrate Judge Wang's September 19, 2025 Order.

[6] The Times also notes that it is not clear what OpenAI seeks to gain from this special 30(b)(6). The Times has provided clear information about the inadvertent deletion of some logs, including answering detailed follow-up questions from OpenAI about ChatExplorer. *See* Dkt. 376-1.

Dated: November 4, 2025

/s/ Ian Crosby

Ian Crosby *(admitted pro hac vice)*
Katherine M. Peaslee *(admitted pro hac vice)*
SUSMAN GODFREY L.L.P.
401 Union Street, Suite 3000
Seattle, WA 98101
Telephone: (206) 516-3880
Facsimile: (206) 516-3883
icrosby@susmangodfrey.com
kpeaslee@susmangodfrey.com

Davida Brook *(admitted pro hac vice)*
Emily K. Cronin (*admitted pro hac vice)*
Adnan Muttalib *(admitted pro hac vice)*
SUSMAN GODFREY L.L.P.
1900 Ave of the Stars, Suite 1400
Los Angeles, CA 90067
Telephone: (310) 789-3100
Facsimile: (310) 789-3150
dbrook@susmangodfrey.com
ecronin@susmangodfrey.com
amuttalib@susmangodfrey.com

Elisha Barron (5036850)
Zachary B. Savage (ZS2668)
Alexander Frawley (5564539)
Eudokia Spanos (5021381)
SUSMAN GODFREY L.L.P.
1301 Avenue of the Americas, 32nd Floor
New York, NY 10019
Telephone: (212) 336-8330
Facsimile: (212) 336-8340
ebarron@susmangodfrey.com
zsavage@susmangodfrey.com
afrawley@susmangodrey.com
espanos@susmangodfrey.com

Steven Lieberman (SL8687)
Jennifer B. Maisel (5096995)
Kristen J. Logan *(admitted pro hac vice)*
ROTHWELL, FIGG, ERNST & MANBECK, P.C.
901 New York Avenue, N.W., Suite 900 East
Washington, DC 20001
Telephone: (202 783-6040
Facsimile: (202) 783 6031
slieberman@rothwellfigg.com

jmaisel@rothwellfigg.com
klogan@rothwellfigg.com

Attorneys for Plaintiff
The New York Times Company