

2025 WL 2691297

Only the Westlaw citation is currently available.
United States District Court, S.D. New York.

IN RE: [OPENAI, INC.](#), COPYRIGHT
INFRINGEMENT LITIGATION,

This Document Relates to:

25-md-3143 (SHS) (OTW), 23-CV-11195

|

Signed September 19, 2025

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

**OPINION & ORDER RE: NYT
CHATEXPLORER LOGS**

[ONA T. WANG](#), United States Magistrate Judge:

*1 Pending now before the Court are OpenAI's and Microsoft's ("Defendants") motions to compel (ECF Nos. 376, 377, 392, 396) ¹ the New York Times Company ("NYT") to produce the content of all chat logs derived from NYT's use of ChatExplorer (the "ChatExplorer Logs"). ² Additionally, OpenAI alleges that because NYT failed to preserve approximately 5 months' worth of ChatExplorer Logs (from November 2023 to March 2024), they seek a limited 4-hour 30(b)(6) deposition regarding NYT's preservation obligations and efforts. (ECF 376). Because Defendants have not demonstrated relevance or proportionality of the discovery sought, Defendants' motions are **DENIED**.

I. BACKGROUND

The Court assumes familiarity with the facts of this case. On September 3, 2024, OpenAI moved to compel NYT to produce, among other things, evidence of NYT's use of nonparties' generative artificial intelligence ("GenAI") tools, as well as documents concerning the NYT's creation and use of its own GenAI products. (*See* 23-CV-11195, ECF 236). The Court denied OpenAI's motion on November 22, 2024, (the "November 2024 Order"), finding the requested discovery irrelevant to OpenAI's fair use defense. *New York Times Co. v. Microsoft Co.*, 757 F. Supp. 3d 594, 599 (S.D.N.Y. 2024) ("This case is about whether Defendant trained their LLMs using Plaintiff's copyrighted material, and whether that use constitutes copyright infringement. It is not a referendum on the benefits of [GenAI], on Plaintiff's business practices, or about whether any of Plaintiff's employees use [GenAI] at work. The broad scope of document production sought here is simply not relevant to Defendant's purported fair use defense."). ³ As part of the November 2024 Order, the Court noted that NYT had agreed to produce documents concerning "the [NYT's] use of the Defendants' [GenAI] tools in reporting or presentation of content, and documents regarding the [NYT's] trainings about Defendants' [GenAI] products." (23-CV-11195, ECF 236 at 1 n. 1). NYT represents that they have produced

2025 WL 2691297

“significant discovery” on NYT’s own use of Defendants’ LLMs, including documents regarding: (1) NYT’s use of Defendants’ LLMs in reporting or presenting content; (2) NYT’s trainings and policies re Defendants’ LLMs (including guidance relating to ChatExplorer); (3) NYT’s AI Initiatives program; and (4) Audience Insight Group reports that assess the impact of GenAI on NYT. (See ECF 475).

*2 Nonetheless, in July 2025, OpenAI and Microsoft both moved to compel production of the content of all of NYT’s ChatExplorer Logs (both prompts and outputs). (ECF Nos. 376, 392).⁴ Additionally, OpenAI alleges that NYT failed to preserve approximately 5 months of ChatExplorer Logs and requests a four-hour Rule 30(b)(6) deposition regarding the NYT’s preservation obligations and efforts. (ECF Nos. 376, 377).⁵ OpenAI and Microsoft assert that the ChatExplorer Logs are relevant to their fair use and substantial non-infringing uses defenses. (ECF Nos. 376, 392). NYT filed oppositions on August 1, 2025, (ECF 417) and August 4, 2025, (ECF 427), arguing the ChatExplorer Logs are not relevant or proportional to the needs of this case.

The Court heard oral argument on August 12, 2025, and directed the parties to submit supplemental briefing by August 22, 2025. (See ECF 459). OpenAI, Microsoft, and the NYT filed their supplemental briefs on August 22. (See ECF Nos. 468, 469, 471, 472, 475, 477).

II. LEGAL STANDARD

Federal Rule of Civil Procedure 26(b)(1) permits discovery of “any nonprivileged matter that is relevant to a party’s claim or defense and proportional to the needs of the case.” The party moving to compel “bears the burden of demonstrating relevance and proportionality.” See *Winfield v. City of New York*, 15-CV-5236 (LTS) (KHP), 2018 WL 840084, at *3 (S.D.N.Y. Feb. 12, 2018); see also *New York Times*, 757 F. Supp. 3d at 597. “When broader discovery is sought, the Court should determine the scope according to the reasonable needs of the action, depending on the circumstances of the case, the nature of the claims and defenses, and the scope of the discovery requested.” *Int’l Code Council, Inc. v. Skidmore, Owings & Merrill, LLP*, 24-MC-412 (DEH) (VF), 2025 WL 1936704, at *2 (S.D.N.Y. July 15, 2025). “It is within a Magistrate Judge’s discretion to decide whether discovery requests are relevant and proportional,” *Radio Music License*

Committee, Inc. v. Broadcast Music, Inc., 347 F.R.D. 269, 273 (S.D.N.Y. 2024), and “Rule 26 gives a district court broad discretion to impose limitations or conditions on discovery[,] which extends to granting or denying motions to compel.” *Edmar Fin. Co., LLC v. Currenex, Inc.*, 347 F.R.D. 641, 646 (S.D.N.Y. 2024).

III. DISCUSSION

I directed the parties to address three principal issues in their supplemental briefing: (1) the relevance of the ChatExplorer Logs; (2) proportionality; and (3) the specifics of burden of production, including the scope of any privilege review. (See ECF 459). For the reasons discussed below, I find that production of the ChatExplorer Logs is not relevant or proportional to the needs of this case.

A. Fair Use

Section 107 of the Copyright Act sets out four non-exclusive factors to consider in determining whether a defendant’s use of a copyrighted work is “fair”:

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

New York Times, 757 F. Supp. 3d at 597. The application of these four factors “requires judicial balancing, depending upon relevant circumstances.” *In re OpenAI, Inc. Copyright Infringement Litigation*, 25-md-3143 (SHS) (OTW), 2025 WL 1652110, at *2 (S.D.N.Y. May 30, 2025) (citing *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 19, 141 S.Ct. 1183, 209 L.Ed.2d 311 (2021)).

*3 Defendants broadly assert that “evidence” of how NYT uses or likely uses OpenAI’s LLMs—for “pro-journalism uses,” (ECF 376 at 4), such as “criticism, comment, news reporting, [] research, ... brainstorming and editing, and

2025 WL 2691297

summarizing Times-owned materials,” (ECF 392 at 3)—is relevant to both the first and fourth fair use factors, such that they are entitled to discovery of the content of the all of the ChatExplorer Logs from NYT employees that used ChatExplorer in their daily work.⁶

1. Fair Use Factor One: The Purpose and Character of the Use

“In assessing the first factor, courts consider two sub-factors: (i) the extent to which the secondary use is transformative and (ii) whether the secondary use is commercial in nature.” *In re OpenAI, Inc. Copyright Infringement Litigation*, 2025 WL 1652110, at *2 (citing *Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 179 (2d Cir. 2024)). OpenAI posits that the content of the ChatExplorer Logs will show that OpenAI’s LLMs are “highly creative and innovative tools” that are used by journalists in ways that advance the overarching goals of copyright law. (ECF 472). Microsoft echoes this sentiment, suggesting that individual employees’ ChatExplorer Logs will show that Defendants’ LLMs are “transformative.” (ECF 392 at 3).

Defendants continue to misconstrue the “use” relevant to the fair use analysis. 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. *New York Times*, 757 F. Supp. 3d at 598-99 (“The fair use factors are concerned with the copier’s use of an original work.”) (internal quotations omitted) (emphasis added).⁷ Indeed, OpenAI themselves cite a case in their supplemental brief supporting this very same principle. (ECF 472 at 3).⁸

The relevant “uses” at issue in this case are: (1) Defendants’ alleged copying of NYT 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.⁹ Defendants do not explain how the content of NYT’s ChatExplorer Logs would have any bearing on how the uses of Defendants’ LLMs might or might not be transformative or commercial. As OpenAI notes in its supplemental brief,

there is no dispute that NYT uses a version of Defendants’ LLMs through ChatExplorer to perform tasks that are not related to producing copyrighted content; the content of the ChatExplorer Logs is not necessary to establish that fact. The ChatExplorer Logs are, at best, duplicative evidence of that fact and are thus not relevant. *See Fed. R. Civ. P. 26(b)(2)(C)* (“[T]he court must limit the frequency or extent of discovery otherwise allowed by these rules or by local rule if it determines that: (i) the discovery sought is unreasonably cumulative or duplicative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive.”).

2. Fair Use Factor Four: The Effect of the Use Upon the Potential Market for or Value of the Copyrighted Work

*4 “In assessing the fourth [fair use] factor, courts ‘focus 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 opt to acquire the copy in preference to the original.’ ” *In re OpenAI, Inc. Copyright Infringement Litigation*, 2025 WL 1652110, at *2 (quoting *Authors Guild v. Google, Inc.*, 804 F.3d 202, 223 (2d Cir. 2015)). Defendants argue that the ChatExplorer Logs are relevant because they “contravene [NYT’s] allegations of market harm” and show that Defendants’ LLMs provide a “public benefit” by allowing the NYT to create more and better content. (ECF 471 at 3-4). Again, Defendants do not explain how the content of the ChatExplorer Logs would support a conclusion of either a public benefit to nonparties or any conclusion of market harm (or lack thereof). And any benefit to NYT from its use of ChatExplorer, even if it could be discovered or inferred from the content of the ChatExplorer Logs, would be probative only of a potential benefit to NYT and would not be probative of any public benefit nor whether Defendants’ LLMs cause market harm.

Even if disclosure of the content of the ChatExplorer Logs yielded the information Defendants contend it would,¹⁰ NYT’s use of ChatExplorer is not relevant to whether other users of Defendants’ LLMs found in those LLMs a “competing substitute” for NYT’s copyrighted works. The fourth fair use factor requires an analysis of the Defendants’ use of NYT’s copyrighted material—not NYT’s use of

2025 WL 2691297

Defendants' LLMs—and whether Defendants' use impacts the market for NYT's copyrighted works.¹¹ The NYT cannot be a “competing substitute” for itself or deprive itself of revenue. *In re OpenAI, Inc. Copyright Infringement Litigation*, 2025 WL 1652110, at *2. OpenAI already has survey results regarding NYT subscription data, and the Court has instructed the parties that questions regarding survey results, subscription data, and analyses of how Defendants' LLMs have “harmed the potential market for or value of” NYT's copyrighted works are better asked at depositions and in expert discovery. (See August 12, 2025, Transcript, ECF 525 at 100); (ECF 459 at 2).

The content of the ChatExplorer Logs is also not relevant to the “public benefits” of Defendants' allegedly infringing uses of NYT copyrighted material. The public benefit consideration of the fourth fair use factor 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.” *Hachette Book Grp.*, 115 F.4th at 195. Private monetary benefits to a copyright holder if a challenged use is permitted is, at best, ancillary to the fair use analysis. *Cf. Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1523 (9th Cir. 1992) (“Public benefit ... may arise because the challenged use serves a public interest. In the case before us, [defendant's use] ... has led to ... [a public] growth in creative expression ... that the Copyright Act was intended to promote.”) (emphasis added) (citing *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 349-50, 111 S.Ct. 1282, 113 L.Ed.2d 358 (1991) (“To this end, copyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work.”) (emphasis added));¹² see also *Nixon v. Source Digital, Inc.*, 23-CV-5218 (JPC), 2024 WL 5202514, at *9 (S.D.N.Y. Dec. 23, 2024) (“Under the fourth factor, courts may also consider any benefit to the general public from the defendant's copying.”) (emphasis added). The ChatExplorer Logs are not probative of NYT's claims of market harm or the alleged public benefit of Defendants' LLMs,¹³ and thus are not relevant to factor four of the fair use analysis under Rule 26.¹⁴

B. Substantial Non-Infringing Uses

*5 Defendants admit that they already have highly relevant evidence in support of their substantial non-infringing uses defense: how Defendants' users are using Defendants' LLMs. (See ECF 469 at 3). However, in conclusory fashion, Defendants argue that they are also entitled to discovery of NYT ChatExplorer Logs because they are “among the strongest evidence supporting” their substantial non-infringing uses defense. (ECF 469 at 3). Microsoft also posits that the ChatExplorer Logs are “uniquely probative of the commercially significant, pro-copyright uses” that *Sony Corp. v. Universal City Studios*, 464 U.S. 417, 442, 104 S.Ct. 774, 78 L.Ed.2d 574 (1984) “cares about” given NYT's commercial presence, and that the NYT actually benefits from its use of Defendants' LLMs. Defendants fail, however, to explain why they believe that NYT's ChatExplorer Logs are “among the strongest evidence” or so “uniquely probative” of the substantial non-infringing uses of their LLMs.

On claims for contributory copyright infringement by material contribution, a plaintiff must establish (1) direct infringement by a third party; (2) a defendant's knowledge of the infringing activity; and (3) that the defendant materially contributed to the third party's infringing activity. *New York Times Co. v. Microsoft*, 777 F. Supp. 3d 283, 305 (S.D.N.Y. 2025). A defendant who otherwise materially contributed to copyright infringement will nevertheless not be liable for contributory infringement if the product is also “widely used for legitimate, unobjectionable purposes or is merely capable of substantial noninfringing use.” *Arista Records LLC v. Lime Group LLC*, 784 F. Supp. 2d 398, 433 (S.D.N.Y. 2011). A substantial non-infringing uses defense requires inquiry into both infringing and non-infringing uses by users of a defendant's product.¹⁵ *Sony Corp. v. Universal City Studios*, 464 U.S. 417, 442, 104 S.Ct. 774, 78 L.Ed.2d 574 (1984).

The content of NYT's ChatExplorer Logs is not relevant to this defense because NYT is the copyright holder and cannot infringe its own copyrights, so every use by NYT is, by definition, non-infringing.¹⁶ The pro-copyright uses that were contemplated in *Sony* necessarily must be uses performed by individuals or entities other than the parties in the case. What is relevant is how other consumers are using Defendants' LLMs, and whether those consumers are using Defendants' LLMs for substantially non-infringing purposes or not. NYT's use of Defendants' LLMs thus cannot be

probative of the existence of substantial non-infringing uses, nor the commercial significance of any such uses.

C. Proportionality Analysis

Fed. R. Civ. P. 26(b) permits discovery of any “nonprivileged matter that is relevant to any party’s claim or defense and proportional to the needs of the case.” Fed. R. Civ. P. 26(b)(1) (emphasis added). “The party seeking discovery has the initial burden of demonstrating relevance and proportionality,” and if they meet that threshold, the “burden then shifts to the party resisting discovery to show undue burden (or a lack of proportionality).” *GMO Gamecenter USA, Inc. v. Whinstone US, Corp.*, 22-CV-5974 (JPC) (KHP), 2025 WL 722853, at *1 (S.D.N.Y. Mar. 5, 2025). Because Defendants have failed to sustain their burden of showing the relevance of the ChatExplorer Logs, the Court need not reach proportionality.

Even if the ChatExplorer Logs were relevant, however, ordering their production would still not be proportional to the needs of this case. Defendants contend, without support, that because (1) the ChatExplorer Logs could hypothetically be exported into a simple spreadsheet of less than 2 gigabytes in file size; (2) attorney client privilege and the reporter’s privilege likely do not apply to the vast majority of the ChatExplorer Logs; and (3) reviewing this spreadsheet for privilege is no more burdensome than reviewing any document for privilege, the burden of production is somehow “minimal.”¹⁷ NYT alleges that they would need to individually review each of the 80,000+ log entries for attorney client privilege and reporter’s privilege. Assuming, conservatively, that it would take 1 minute to review each individual entry, it would take approximately 1,333 hours to review the ChatExplorer Logs in full. If NYT assigned two human attorneys, each billing at \$750

per hour 60 hours a week to conduct this review, it would still take 11 weeks (nearly 3 months) and cost upwards of \$1 million for the privilege review alone. Defendants do not carry their burden to show proportionality, relying as they have on conclusory statements that the information sought is “uniquely probative” and that the burden of production would be “minimal.”¹⁸

D. OpenAI’s Spoliation Allegations

*6 Finally, the Court addresses OpenAI’s request for a special 4-hour Rule 30(b)(6) deposition regarding NYT’s preservation obligations and efforts with respect to the ChatExplorer Logs. (ECF 376 at 4). Because the ChatExplorer Logs are not relevant, the inadvertent loss¹⁹ of these logs does not implicate Fed. R. Civ. P. 37(e). *See In re Keurig Green Mountain Single-Serve Coffee Antitrust Litigation*, 341 F.R.D. 474, 497 (S.D.N.Y. 2022) (explaining 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”) (internal quotation marks omitted).

IV. CONCLUSION

For the foregoing reasons, Defendants’ motions to compel production of the ChatExplorer Logs and for a special 30(b)(6) deposition are **DENIED**.

SO ORDERED.

All Citations

--- F.Supp.3d ----, 2025 WL 2691297

Footnotes

- 1 Unless otherwise indicated, all references to the docket refer to 25-md-3143.
- 2 ChatExplorer is an internal tool developed and used by NYT to interface with OpenAI’s GPT models using OpenAI’s API. (*See* ECF Nos. 392, 417). The ChatExplorer Logs are the actual content underlying NYT employees’ use of ChatExplorer, i.e., employees’ prompts to ChatExplorer and the outputs they received

from Defendants' LLMs. NYT represents that after a brief period of limited internal use, ChatExplorer became generally available to NYT employees on December 13, 2023, (ECF 417 at 1), two weeks before NYT filed the complaint in Case No. 23-CV-11195 (see 23-CV-11195, ECF 1). Ostensibly, NYT employees still use ChatExplorer today. (See August 12, 2025, Transcript, ECF 525 at 15-16) ("Indeed, I think I checked before this hearing, and just in the last two months there are an additional 20,000 log entries that we've captured and preserved.").

- 3 Defendants objected to the November 2024 Order in December 2024. (See 23-CV-11195, ECF Nos. 362, 368). In May 2025, Judge Stein denied Defendants' Rule 72(a) motions to set aside the November 2024 Order. (ECF 82).
- 4 The public, redacted version of Microsoft's motion is available at ECF 392. The unredacted version of the motion remains under temporary seal at ECF 396 until Microsoft's related sealing motion is resolved. (See ECF 387).
- 5 The public, redacted version of OpenAI's motion is available at ECF 376. The unredacted version of the motion remains under temporary seal at ECF 377 until OpenAI's related sealing motion is resolved. (See ECF 373).
- 6 Microsoft and NYT dispute whether NYT previously agreed to produce any of the ChatExplorer Logs. (See ECF Nos. 392, 427).
- 7 See also [Hachette Book Grp.](#), 115 F.4th at 180-181 (addressing defendant's use as follows: "We conclude that IA's use of the Works is not transformative. IA creates digital copies of the Works and distributes these copies to its users in full, for free. Its digital copies do not provide criticism, commentary, or information about the originals. Nor do they add something new, with a further purpose or different character, altering the originals with new expression, meaning or message ... IA argues that ... its use of the Works is transformative because of the sort of lending the scanning enables, namely, (1) making lending more efficient, and (2) enabling other uses not possible with print books and physical borrowing. Neither of these justifications render IA's use of the Works transformative because the underlying purpose of making the Works available in a derivative format is still the same.") (emphasis added) (internal citations and quotation marks omitted).
- 8 "These uses show that OpenAI's LLMs 'add something new, with a further purpose' to the asserted works. See also [Campbell v. Acuff-Rose Music, Inc.](#), 510 U.S. 569, 579, 114 S.Ct. 1164, 127 L.Ed.2d 500 (1994) (the 'central purpose' of factor one is determine whether the defendant's use is 'transformative' in that it 'adds something new, with a further purpose.')." (ECF 472 at 3) (emphasis added).
- 9 In a similar case in the Northern District of California, Judge Alsup's analysis of the defendant's fair use defense was confined to the defendant's use of the plaintiff's copyrighted materials. See [Bartz v. Anthropic PBC](#), — F. Supp. 3d —, —, 2025 WL 1741691, at *7 (N.D. Cal. 2025) ("All agree that one use at issue was [Defendant] training LLMs to receive text inputs and return text outputs. More specifically, Anthropic used copies of Authors' copyrighted works to iteratively map statistical relationships between every text-fragment and every sequence of text-fragments so that a completed LLM could receive new text inputs and return new text outputs as if it were a human reading prompts and writing responses.") (emphasis added).
- 10 Defendants argue that ChatExplorer Logs will show "countless other potential examples" of how NYT uses Defendants' LLMs, including analyzing subscription data, consumer insights, and/or financial records; translating works; developing new recipes; generating captions for photos; and more. (ECF 472 at 3).

- 11 See [Bartz](#), — F. Supp. 3d at —, 2025 WL 1741691, at *1 (“The copies used to build a central library and that were obtained from pirated sources plainly displaced demand for Authors’ books – copy for copy.”) (emphasis added).
- 12 [Sega](#) does not stand for the proposition that a positive financial effect on the plaintiff is “particularly compelling.” (ECF 472 at 4). While the [Sega](#) court acknowledged that Sega may have derived some financial benefit from the defendant’s copying, the court still focused on the “public benefit” of increased “dissemination of other creative works.” [Sega](#), 977 F.2d at 1523.
- 13 Even assuming *arguendo* that NYT’s use of Defendants’ LLMs for other purposes besides journalism is at all probative of a “public benefit,” Defendants fail to explain why they need all of NYT’s employees’ ChatExplorer Logs. As discussed above, the fact that the NYT uses Defendants’ LLMs for other such purposes is relevant, and NYT represents that it has already produced substantial discovery on that fact. (See ECF 427 at 2).
- 14 Defendants advance two additional arguments seemingly unrelated to the fair use factors. First, Defendants suggest that ChatExplorer Logs should be produced because OpenAI is “entitled to any evidence of failed efforts at regurgitation from Times employees.” (ECF 376 at 4). However, OpenAI does not even allege that there is any reason to believe that an internal tool rolled out to NYT employees after this lawsuit’s inception (ECF 417 at 1) and seemingly used in the ordinary course of business would contain evidence of NYT employees trying and failing to regurgitate their own work.
- Second, Defendants cite [U.S. v. Arroyo](#), 600 F. App’x 11 (2d Cir. 2015)—a non-precedential summary order in a criminal case—for the proposition that the ChatExplorer Logs are “especially relevant” to rebut a “central trial theme” of NYT. While the legal proposition—evidence that rebuts an opponent’s central trial theme may be “especially relevant”—is not incorrect, OpenAI has not shown how the content of ChatExplorer Logs supports or rebuts NYT’s claim of market harm. (ECF 472 at 4). OpenAI’s argument here seems to be that if any NYT employee uses Defendants’ LLMs at work, their work is “pro-journalism” because their employer’s work is independent journalism, (ECF 72 at 1), and thus the ChatExplorer Logs would tend to show that Defendants’ LLMs actually “support journalism” rather than causing market harm. If that is the argument, however, then only the fact of the use—not the individual prompts and outputs—is necessary. If that is not the argument, then OpenAI is asserting that the content itself matters, and one embedded assumption is that the responses are necessarily material, truthful, and reliable to some degree.
- 15 See, e.g., [Arista Records](#), 784 F. Supp. 2d at 434 (noting that while most users used LimeWire “overwhelmingly for infringement,” others used it to share noninfringing content such as books and movies in the public domain).
- 16 Even if NYT’s use was at all probative, the de facto non-infringing nature of NYT’s use would render the individual ChatExplorer Logs duplicative.
- 17 Moreover, Defendants apparently seek all ChatExplorer Logs that exist, which also weighs against a proportionality finding.
- 18 Nor has NYT “put at issue” its use of Defendants’ LLMs such that they have waived any privilege over the content of the ChatExplorer Logs. As explained above, the “end uses” at issue in this case are the Defendants’ use of Plaintiffs’ copyrighted works and those of other commercial users of Defendants’ LLMs, not NYT’s use of ChatExplorer.

2025 WL 2691297

19 NYT represents that after in-house counsel directed a litigation hold be placed on the ChatExplorer Logs, a NYT employee tasked with implementing the litigation hold inadvertently “believed those logs were already under a one-year preservation policy.” (ECF 417). However, the ChatExplorer Logs were stored in a “SumoLogic partition” with a 90-day retention period. (*Id.*). After this error was discovered, NYT rectified the issue and migrated the ChatExplorer Logs to a separate partition with a 360-day retention policy and simultaneously wrote the data to “Amazon S3 with a 10-year retention policy.” (*Id.*).

End of Document

© 2025 Thomson Reuters. No claim to original U.S. Government Works.