SUSMAN GODFREY L.L.P.

March 5, 2026

Hon. Ona T. Wang
Southern District of New York
500 Pearl Street
New York, New York 10007

Re: *In re OpenAI Copyright Infringement Litigation*, No. 1:25-md-03143
This Document Relates to: *NYT v. Microsoft et al.*, No. 1:23-cv-11195
Ziff Davis, Inc., et al. v. OpenAI, Inc., No. 1:25-cv-04315
Opposition to Defendants' Motions to Compel, ECF Nos. 1370 & 1377

Dear Magistrate Judge Wang:

Plaintiffs The New York Times Company ("The Times") and Ziff Davis oppose OpenAI's motion to compel. *See* Dkt. 1377. The Times further opposes Microsoft's motion to compel. *See* Dkt. 1370.

I. OpenAI's Motion to Compel[1]

A. The New York Times

1. OpenAI's Request for Further BrandMatch Documents Seeks Material that This Court Has Already Held to be Irrelevant.

First, OpenAI seeks to compel the production of further documents related to The Times's BrandMatch tool. Dkt. 1377 at 2. Specifically, OpenAI claims that "BrandMatch reflects The Times's use of OpenAI's technology to develop a product that generates profits for The Times," which is allegedly relevant to the fourth fair use factor as it shows how "The Times itself is monetizing the technology." *Id.* The Court has already rejected this theory of relevance.

As this Court has previously noted, "[t]he 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." Dkt. 604 at 8 (emphasis in original) (quoting *Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 178 (2d Cir. 2024)). "Private monetary benefits to a copyright holder if a challenged use is permitted is, at best, ancillary to the fair use analysis." *Id.* (emphasis in original). Accordingly, any purported private benefit The Times derived from its development of the BrandMatch tool (for which it

[1] At the outset of its letter motion, OpenAI proclaims that "Plaintiffs have had full access to the evidence necessary to litigate their claims, and fact discovery has been more than sufficient to permit the case to proceed to expert discovery." Dkt. 1377 at 2. Plaintiffs will not rehash every aspect of the significant deficiencies laid out in their affirmative motion pending before the Court. *See* Dkt. 1383. However, as noted in Plaintiffs' letter to the Court, OpenAI has ***still*** "failed to produce readily available, searchable, and already de-identified logs of ChatGPT conversations." Dkt. 1384 at 1. Output logs are among the most relevant evidence in this case, bearing "directly on how OpenAI products may generate outputs that infringe, substitute for, or dilute the markets for Plaintiffs' copyrighted works." *Id.* OpenAI's failure to produce such pivotal evidence in a readily accessible manner has clearly prejudiced Plaintiffs in their ability to prosecute their action and develop their case.

appropriately paid and licensed OpenAI's models) is not probative of "the alleged public benefit of Defendants' LLMs, and thus [documents related to BrandMatch] are not relevant to factor four of the fair use analysis under Rule 26." *Id.*

Furthermore, OpenAI's contention that the "production of BrandMatch documents is consistent with The Times's production of its ChatExplorer logs after the recent Rule 72 argument in front of Judge Stein" is completely disingenuous. Dkt. 1377 at 2. Judge Stein never opined on whether documents relating to The Times's use of OpenAI's products were relevant, nor did he alter this Court's ChatExplorer ruling in any manner. *See* Dkt. 604. [redacted] As this Court previously acknowledged, this case is not about The Times's use of Generative AI tools. *See* NYT Dkt. 344; NYT Dkt. 592. That has not changed.

In any event, The Times has already produced dozens of documents regarding BrandMatch. For instance, NYT_00269257 provides a description of BrandMatch, the LLMs leveraged by the tool, and the launch date. *See* Ex. B (NYT_00269257) [redacted] *see also* NYT_01059228; NYT_01059268; NYT_01059352. OpenAI has also asked multiple Times employees—including Zach Seward, Bailey Evans, and Joy Robins—about BrandMatch. *See, e.g.*, Bailey Evans Dep. Tr. 183:3–15 [redacted]. OpenAI is not entitled to more.

2. <u>The Court Should Deny OpenAI's Request for Copies of The Times's "Grids."</u>

OpenAI's recounting of supposedly withheld "relevant financial documents" misframes what these documents actually are: namely, internal working tools that are neither the underlying data source nor the ultimate output for any Times forecast or budget. Particularly in light of The Times's fulsome financial productions in this case–which contain The Times's fully verified financial data–OpenAI's request for these internal working documents should be denied.

As an initial matter, OpenAI has never, whether in meet and confers or in their opening brief, identified ***any*** specific information they need from these working grids, despite the fact that Defendants have had exemplar copies of grids for over a year (and, given that OpenAI has long had examples of this document, their claim that it was "recently revealed" (OAI Br. at 2) is simply false). Ex. C (February 24, 2026 Email from K. Peaslee to A. Espeland).[2]

[2] In the parties' correspondence, OpenAI demanded not only the Grid documents themselves, but also a highly burdensome production of documents and communications related to the Grid. Ex. C. The Times reserves the right to raise further objections should OpenAI improperly attempt to revive those abandoned demands.

Furthermore, as The Times has repeatedly explained to OpenAI, the "Grid" spreadsheets highlighted by OpenAI are internal working documents used by The Times's finance team as an organizing tool. They are not "the budget," and OpenAI's citation to Ms. Yang's general description of them does not prove otherwise—Ms. Yang was not asked at her deposition whether this was a formal or final document; this spreadsheet fell well outside the scope of her designated 30(b)(6) topics; and The Times employee more familiar with these documents made clear that the Grid was simply a "spreadsheet output" and was not a "source" for The Times's budgets or forecasts. Gurian-Peck Tr. at 56:20-57:7. OpenAI knows this: The Times has separately produced actual budgeting materials and has emailed about them with Defendants. *See* NYT_00871921 (September 2023 Board of Directors Meeting Presentation); NYT_01034150; Ex. D (February 23, 2026 Email from K. Peaslee to A. Schreiber). Indeed, The Times has produced budgeting and forecasting spreadsheets from 2023 and 2025 that are far more detailed than these working grids—and, at Defendants' request, has even produced the dozens of underlying models that populate those spreadsheets. *See* Ex. D. OpenAI has not, and cannot, explain why it needs more. Accordingly, the Court should deny OpenAI's request for yet more documents that it does not need.

3. <u>The Times Has Already Produced the Relevant Referral Traffic and Conversion Data.</u>

OpenAI next "requests that the Court order The Times to produce data it has collected regarding referral traffic from AI products and search, as well as data regarding the registrations and subscriptions generated from traffic referred by these sources." Dkt. 1377 at 3. OpenAI further states that "The Times has refused to confirm it has produced all of the referral traffic data it maintains for AI products and search." *Id.*

As an initial matter, The Times notes that OpenAI made no attempt to meet and confer on these documents or to clarify the scope of its requests. Instead, on March 2, 2026 at 2:56pm—three days after the close of fact discovery—The Times received an email, for the first time, requesting these documents. Ex. E (March 2, 2026 Email from A. Espeland to A. Muttalib). OpenAI has known about this Generative AI traffic spreadsheet for months, and even asked Hannah Poferl about this document at her December 5, 2025 deposition. *See* Ex. F (Hannah Poferl 30(b)(6) Dep. Tr.) at 132:16–135:16. [redacted] Yet, despite Ms. Poferl noting this fact over four months ago, OpenAI only followed up on this information ***after*** the close of fact discovery. Consistent with the Court's repeated guidance, OpenAI's untimely request and its failure to meet and confer on this request is inappropriate and is an independent ground to deny its request.

Even if the Court is willing to consider OpenAI's request, OpenAI has not and cannot establish relevance of OpenAI's request for referral traffic from, for instance, Anthropic or Perplexity. *See* Dkt. 344 (holding that documents related to "nonparties' GenAI tools are neither relevant nor proportional to the needs of the case"); Dkt. 592 (same). And, to the extent OpenAI's request is limited to referral traffic from Defendants' own products, The Times has already

produced such evidence. Specifically, The Times has already produced NYT_01030093, which provides weekly traffic data from ChatGPT to The Times from October 31, 2022 to November 3, 2025. *See* Ex. G (Generative AI Traffic Metrics). In other words, this Excel spreadsheet provides the weekly referral traffic data from ChatGPT to The Times from the date the product originated. The Times has also produced data showing its Bing search performance over the last 24 months, and its Google search performance over the last 16 months. *See* NYT_01072773; NYT_01072774. Specifically, this data shows on a daily basis, *inter alia*, the number of users who searched for The Times's content, the total number of clicks to The Times's website, and the associated click-through rate. Finally, The Times has also produced data from January 1, 2023 to the present showing the number of weekly referrals sent from search platforms to The Times's webpages. *See* NYT_01072768. The Times is unaware of what further information OpenAI is seeking.

B. Ziff Davis

1. <u>OpenAI's Motion to Compel Ziff Davis to Produce Additional Documents is Based on Mischaracterizations of the Record and Should be Denied.</u>

OpenAI's belated demand for a wide range of additional documents related to Ziff Davis's use of OpenAI's products rests on two fundamental distortions of the record. First, OpenAI asserts that Ziff Davis's deposition testimony establishes [redacted] Second, OpenAI claims that [redacted]. Neither assertion is true, and there is simply no support for OpenAI's position.

As a fundamental matter, this Court has repeatedly made clear that the use of generative AI products by plaintiffs is either irrelevant or tangential at best. *See* Docket Nos. 604, 704. Ziff Davis nevertheless has produced a significant number of documents [redacted]

In any event, the deposition testimony [redacted] To the extent any testimony touches on the use of OpenAI products, [redacted] While several deponents did refer to [redacted] *See* Ex. H (Excerpts of Mattheis 30(b)(6) Depo) at 15:10-17 [redacted] Characterizing any of this as [redacted] is wholly unsupported by the record.

Indeed, Ziff Davis witnesses repeatedly and consistently testified that [redacted] *See e.g.*, Ex. I (Excerpts of Oh 30(b)(6)

Day 1 Rough Tr.) at 191:15-20 [redacted]; Ex. J (Excerpts of Mattheis Depo) at 76:21-22 [redacted] Ex. H (Excerpts of Mattheis 30(b)(6) Depo) at 15:10-17 [redacted]; Ex. K (Excerpts of Turrentine 30(b)(6) Depo) at 14:17-25 [redacted] Ex. L (Excerpts of Schneider 30(b)(6) Depo) at 23:13-14 [redacted]; *id.* at 28:8-14 [redacted]).

As for OpenAI's argument that [redacted] even putting aside that OpenAI has never raised this argument at any time until now, this too simply not true. In fact, Ziff Davis's witnesses testified [redacted]. In particular, contrary to OpenAI's assertions, Ziff Davis's corporate representative expressly testified that [redacted] *See* Ex. M (Excerpts of Oh 30(b)(6) Day 2 Rough Tr.) at 35:2-16; *see also* Ex. N (Excerpts of Fortuna 30(b)(6) Rough Tr.) at 73:18-76:15 [redacted]. If OpenAI wants to establish a license defense, it bears the burden of doing so, e.g., with its own records. It cannot simply misrepresent or ignore clear testimony [redacted], and attempt to shift the burden to Ziff Davis to prove that no license exists.

In light of all of this, OpenAI's sweeping, last-minute demands for additional documents are entirely unsupported by the record and rest on false characterizations of the deposition testimony and should be rejected.

First, OpenAI seeks records [redacted] But these records should already be in OpenAI's possession. As the vendor, OpenAI presumably maintains account records, billing records, and user data for every account associated with its customers, including Ziff Davis. Compelling Ziff Davis to reconstruct records that OpenAI already possesses — and that OpenAI is better positioned

to produce accurately — would impose disproportionate burden for no informational gain, especially given the lack of any evidence or testimony supporting the contention that [redacted].

Second, Ziff Davis has already produced [redacted] If OpenAI contends otherwise, it should identify with specificity what it believes is missing.

Finally, OpenAI seeks documents [redacted]. But OpenAI has not made the predicate showing that would justify this wide-ranging and intrusive discovery. As noted above, the deposition record contains no testimony [redacted]. Nor do any produced documents suggest this. [redacted]. OpenAI cannot bootstrap these marginal uses into broad discovery of every employee's AI interactions on the theory that a license defense might materialize.

For all of these reasons, OpenAI's motion should be denied.

II. Microsoft's Motion to Compel

1. The Times Has Already Produced All Relevant Subscription Metrics.

Microsoft first requests that The Times produce "[a]n export of monthly data from 2015 to 2025 covering The Times's: (1) overall anonymous, registered, and subscribed users; (2) paywall conversion rates for anonymous and registered users; (3) weekly and monthly active users broken out by these three groups; (4) conversion and registration rates for anonymous and registered users; (5) overall churn and cancellation rates; and (6) conversion rates from Defendants' GenAI tools referral traffic." Dkt. 1370 at 2. Microsoft's request is untimely and overbroad, and should be denied.

Microsoft raised its request for vast amounts of data, for the first time, at 10:06 pm on February 27, 2026—less than ***two hours before the close of fact discovery***. *See* Ex. O (February 27, 2026 email from A. Schreiber to A. Muttalib). Microsoft's claim that it could only seek this data after completing Hannah Yang's deposition earlier that day is entirely false. As Microsoft itself concedes, it has had access since 2024 to "Board presentations that reflect anonymous and registered users' conversion rates." Dkt. 1370 at 3. Indeed, Microsoft asked about the existence of this very data at Hannah Poferl's December 5, 2025 deposition. *See* Ex. F (Hannah Poferl 30(b)(6) Dep. Tr.) at 45:14–46:15 [redacted]. Microsoft has similarly had documents showing that The Times tracks weekly active users broken down by anonymous, registered, and subscribed users since at least May 2025. *See, e.g.*, NYT_00841145 (document produced in May 2025 providing detailed breakdown of anon/regi/subs for Games and Cooking on a weekly basis). And the fact that The Times tracks churn and cancellations is apparent on the face of numerous documents and reports

from its Audience team. Yet, despite this, Microsoft only reached out for this data long after business hours on the last day of discovery. Given Microsoft's (literal) 11th hour ask, the parties have not met and conferred regarding this newly requested data. Consistent with the Court's repeated guidance, Microsoft's untimeliness and its failure to meet and confer on this request is an independent ground to deny its request.

Even if the Court is willing to entertain Microsoft's request, The Times has already satisfied its production obligations in response to Microsoft's RFPs. Specifically, Microsoft's RFPs request:

> · Documents ***sufficient to show the number of Your active subscribers*** for each month beginning January 1, 2018 to the present (Microsoft RFP 59);
>
> · Documents concerning increases, decreases, fluctuations, or changes in subscription rates to The New York Times (including print and digital editions) since January 1, 2019, including any study or analysis of the causes underlying same (Microsoft RFP 74), to which The Times responded it would produce documents ***sufficient to show*** trends in subscription rates to The New York Times print and digital editions since January 1, 2019 (a limitation on Microsoft's overbroad RFP that it has never contested); and
>
> · Documents ***sufficient to show*** the rate at which Your readers who encounter a paywall convert to a trial or digital subscription of Your publication, or otherwise become registered users from January 1, 2015 to present (Microsoft RFP 143).

Consistent with these requests, The Times has produced documents sufficient to show the number of its active subscribers from January 1, 2018 to the present, as well as documents concerning increases, decreases, fluctuations and changes in subscription rates to The Times. For instance, The Times produced audience reports and monthly Flashbooks, which contain The Times's total number of monthly subscribers and its subscription revenue broken down by product. Furthermore, The Times has produced multiple data reports, which contain the number of weekly subscription starts (by product) as well as paywall conversion rates. *See* NYT_01036107; NYT_01072768. The Times has also already produced data regarding its overall churn and cancellation rates. *See* NYT_00941696; NYT_01072768. And, The Times does not maintain "conversion rates from Defendants' GenAI tools referral traffic."[3] Dkt. 1370 at 2. Accordingly, The Times has satisfied its discovery obligations with respect to Microsoft's "sufficient to show" requests.

Microsoft's new request for The Times to now ***also*** produce various datasets related to The Times's anonymous, registered, and subscribed users is not responsive to any of its "sufficient to show" requests. And Microsoft's sole basis for seeking this discrete information (i.e., subscription data broken down by these three specific categories) is that Hannah Yang allegedly testified that

[3] To state the obvious, The Times's Audience team cannot track the key metric at issue in this case – the users The Times cannot convert into subscribers because Defendants' tools do not refer those users to The Times's website – because that data is solely in Defendants' possession.

such data is needed for an "apples-to-apples comparison of The Times's subscription conversion rate (i.e., the rate at which anonymous or registered users convert to become subscribers)." Dkt. 1370 at 2. That mischaracterizes Ms. Yang's testimony. In making that statement, Ms. Yang was noting that to compare the conversion rate over two distinct periods of time, the relevant timeframe between the two data sets needs to be identical (i.e., weekly data in 2021 vs. weekly data in 2025). *See* Hannah Yang 30(b)(6) Dep. Tr. at 136:17-137:3 ("Q: And what information or component would you say you need to make that determination? A: I would like to know whether the audience number is weekly on the other 2021 audience number."). Ms. Yang was **not** opining on this discrete dataset that Microsoft claims to seek. Microsoft otherwise does not justify its need for this discrete information, especially in light of The Times's extensive production regarding its overall subscription numbers. This is particularly true given that fact discovery has already concluded.

Furthermore, Microsoft's request for a 30(b)(6) witness "to testify regarding subscriber data metrics" is a nonstarter. Microsoft had access to The Times's voluminous subscription data metrics well before Hannah Poferl and Hannah Yang's respective 30(b)(6) depositions. Any failure to ask questions about this data is solely attributable to Defendants.

The Court should deny Microsoft's untimely and unjustified request for voluminous data (i.e., monthly data over the last ten years for four discrete categories of information).

2. **The Times Has Produced Extensive Information Regarding its Brand Tracker Surveys.**

Microsoft next asks the Court to "compel The Times to produce the survey instruments and underlying methodology for the surveys it intends to rely on." Dkt. 1370 at 4. Specifically, Microsoft claims "The Times cannot evade the requirements of FRE 702 and 703 by conducting new 'internal' surveys while litigation is pending and then refusing to subject those surveys to any meaningful scrutiny." *Id.* As an initial matter, The Times notes that these surveys were performed in the normal course of The Times's business and not conducted in anticipation of litigation (were it otherwise, they would be privileged). Indeed, The Times began conducting these surveys before this litigation commenced. *See, e.g.*, NYT_01019606 (User Feedback Sprint 207 AIG Report, dated August 2023). As such, FRE 702 and 703 are not applicable to this request.[4] Furthermore, Microsoft has long been aware of these surveys. Its last-minute request for new survey data continues Defendants' trend of making unreasonably burdensome requests.

Nevertheless, in an effort to resolve this dispute, The Times is willing to produce the requested survey instruments and methodology associated with its survey data. However, Microsoft's request for additional 30(b)(6) time with Sonia Yamada is baseless and should be denied.

Microsoft claims it is entitled to additional deposition time because "after Defendants deposed Ms. Yamada, The Times belatedly produced" brand tracker survey data that was allegedly

[4] To the extent Microsoft's argument is that The Times is required to produce these surveys because its experts will rely on them, that argument is premature. There have been no expert reports submitted in this matter and, as such, this challenge is not ripe for the Court's consideration.

relevant to her topics. Dkt. 1370 at 4. That is inaccurate. As an initial matter, Defendants did not serve a single topic on the Times' brand tracker reports or their underlying data. The Times only agreed to have Ms. Yamada testify to brand tracker reports themselves, and did so in an effort to avoid a dispute and move discovery along. Furthermore, all brand tracker surveys were produced well before Ms. Yamada's deposition. The only data that was produced after Ms. Yamada's deposition were qualitative GenAI studies, which provide transcript excerpts of interviews with NYT readers. This data did not concern The Times's brand tracker surveys, which are ***quantitative*** tools. In short, Ms. Yamada was not designated to testify on the additionally produced studies (i.e., studies unrelated to brand tracker), and Defendants should not be able to seek additional 30(b)(6) time on this basis.

Finally, Ms. Yamada testified in great detail on her designated topics—*i.e.*, The Times's subscription cancellation surveys, *see, e.g.*, Yamada Tr. 15:22–31:19, 34:25–35:23, 44:18–49:15, brand tracker reports, *see id.* 49:23–53:3, and the underlying brand tracker surveys, *see id.* 56:2–70:20. And, Defendants failed to make any request on the record to keep the deposition open. Thus, to the extent Defendants claim that Ms. Yamada was unprepared, that is expressly contradicted by the record.

Respectfully submitted,

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

/s/ Steven Lieberman
Steven Lieberman
Rothwell, Figg, Ernst & Manbeck

/s/ Lacy H. Koonce, III
Lacy H. ("Lance") Koonce, III
Klaris Law PLLC

cc: All Counsel of Record (via ECF)