SUSMAN GODFREY L.L.P.

July 3, 2024

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

> Re: *The New York Times Company v. Microsoft Corporation, et al.*,
> Case No.: 23-cv-11195-SHS: Opposition to OpenAI's Motion to Compel

Dear Judge Stein:

Plaintiff The New York Times Company ("The Times") opposes OpenAI's July 1 motion seeking to compel discovery into, among other things, the newsgathering process for every Times article copied by OpenAI. Dkt. 152. OpenAI's motion should be denied in full.

1. OpenAI's Demand for Confidential Reporters' Files is Irrelevant, Improper, and Harassing (RFPs 10-12).

OpenAI's claim that it needs all "reporter's notes, interview memos, records of materials cited, or other 'files' for each asserted work"—purportedly to determine whether The Times's works are in fact protectable intellectual property—is unprecedented and turns copyright law on its head. OpenAI cites no caselaw permitting such invasive discovery, and for good reason. It is far outside the scope of what's allowed under the Federal Rules and serves no purpose other than harassment and retaliation for The Times's decision to file this lawsuit.

First, The Times's newsgathering process on a story-by-story basis has no relevance to whether it is entitled to enforce the millions of copyrights it has registered over the years. OpenAI claims that the reporters' notes underlying the asserted works may shed light on whether The Times's news articles are really original, expressive content—but that is not how copyright law works. The expressive nature of a work is determined by reference to the work itself. *See Salinger v. Random House, Inc.*, 811 F.2d 90, 98 (2d. Cir. 1987) (reiterating that copyright law protects the "manner of expression" in an author's work, including structure, word choice, and "the author's analysis or interpretation of events"). Moreover, even in the improbable case that a reporter's notes show that 90% of an article comprises verbatim quotes from the author's original sources, that article would still be protected by copyright. *See* 17 U.S.C. § 101 (a "work based upon one or more preexisting works" is a protected derivative work); *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., Inc.,* 499 U.S. 340, 348-49 (1991) (affirming that copyright protection extends to original components of compilations of facts arranged independently by an author); *Wozniak v. Warner Bros. Ent. Inc.*, 2024 WL 1311856, at *15 (S.D.N.Y. Mar. 27, 2024) (applying copyright protection to material contributed by the author of a derivative work). The Times, as copyright holder, would thus be able to sue for infringement of such an article—particularly when (as here) Defendants appear to have copied entire works verbatim.

July 3, 2024
Page 2

Notably, OpenAI does not cite a single case in which a news organization has been ordered to produce such materials. OpenAI is not entitled to unbounded discovery into nearly 100 years of underlying reporters' files, on the off chance that such a frolic might conceivably raise a doubt about the validity of The Times's registered copyrights.[1] *See* Fed. R. Civ. P. 26(b)(1) (discovery must be relevant to a party's claim or defense and proportional to the needs of the case). OpenAI has the burden to show that the discovery it seeks "is more than merely a fishing expedition." *Walsh v. Top Notch Home Designs Corp.*, 2022 WL 3300190, at *3 (E.D.N.Y. Aug. 11, 2022) (quoting *Evans v. Calise*, 1994 WL 185696, at *1 (S.D.N.Y. May 12, 1994)). It has not done so.

Second, even if there could potentially be something relevant within the vast universe of documents OpenAI demands, its overbroad discovery requests are neither necessary nor appropriate. The Times has agreed to produce the works themselves, and its certificates of copyright registration for those works, which constitute prima facie evidence of validity of the copyright and of the facts stated in the certificate. *See* 17 U.S.C. § 410(c). And with respect to The Times's investment in its journalism, Dkt. 152 at 2, The Times has agreed to produce documents that show its financial investment into the creation of its published works (RFP 32).

Third, this discovery is improper and seeks to invade the reporters' privilege.[2] Permitting OpenAI to investigate The Times's privileged newsgathering process would have serious negative and far-reaching consequences. It would entail the disclosure of The Times's confidential reporters' files on investigative reporting into highly sensitive matters, including those related to the Defendants themselves. OpenAI's request imposes an undue burden on The Times that is disproportionate to (and indeed, wholly divorced from) the needs of this case. Fed. R. Civ. P. 26(b)(1).

That OpenAI claims for the first time in this motion not to be seeking the identities of confidential sources makes little difference. The burden of separating information from which identities of confidential sources could be derived from the requested materials would be enormous, if it could even be done. In any event, reporters' source files as a whole are protected by the reporters' privilege, regardless of whether such files reveal the identities of confidential sources. *See Gonzales v. Nat'l Broad. Co., Inc.*, 194 F.3d 29, 33-34, 36 (2d Cir. 1999) (recognizing qualified privilege under federal law for confidential and non-confidential journalistic source materials). OpenAI does not dispute that these materials are protected by federal law, nor does it begin to address the chilling effect that such massive discovery requests would have on a news organization's reporting—and its ability to bring lawsuits to defend its copyrighted works. Indeed, given the wildly improper scope of this request, one has to wonder if a chilling effect is exactly

---

[1] OpenAI asserts that such discovery might show that Times articles are themselves "authored by artificial intelligence." Dkt. 152 at 2. OpenAI gives no basis for this wild speculation, which is particularly absurd because the vast majority of Times Works were written before the development of consumer generative AI products.

[2] Contrary to OpenAI's assertion, the federal reporters' privilege and New York Shield Law both apply because The Times asserted claims under federal and New York law. *See* Compl. ¶¶ 192-197 (common law unfair competition claim).

July 3, 2024
Page 3

what OpenAI, who appears to have stolen from millions of content creators, is hoping for. The Court should reject OpenAI's effort to weaponize discovery as a means of targeting the confidential and irrelevant information that underpins The Times's reporting. The Times's newsgathering process is not on trial. OpenAI and Microsoft's infringement of millions of The Times's registered copyright works is.

2. Unsubstantiated Allegations of Infringement Are Irrelevant (RFPs 8, 9, & 13).

OpenAI's RFPs 8, 9, and 13 seek documents and communications relating to unadjudicated allegations of infringement or plagiarism against The Times, arguing—without support—that mere allegations of infringement or plagiarism are relevant to rebut the validity of The Times's copyrights. The Times agreed to produce quasi-judicial and judicial determinations that any of the asserted works infringed a third party's rights, if any exist. Dkt. 152 at 3. This agreement covers the only potentially relevant information in this category, which are determinations that The Times actually infringed a third party's work. Anything more is irrelevant and undiscoverable. Fed. R. Civ. P. 26(b)(1). Any side inquiry into the validity of third-party complaints against The Times would waste the Court's time on mini trials and detract from the real issues—OpenAI and Microsoft's *actual* infringement of The Times's copyrights (not The Times's hypothetical infringement of someone else's).[3]

3. Communications with the USCO are Irrelevant (RFP 14).

OpenAI's request for all of The Times's communications with the U.S. Copyright Office ("USCO") should also be denied. In addition to any deposit copies The Times has in its possession, custody, or control, The Times agreed to produce copyright registration certificates issued by the USCO. This is prima facie evidence of validity of the copyright and of the facts stated in the certificate. 17 U.S.C. § 410(c). OpenAI has not raised any defense calling The Times's communications with the USCO into question. Correspondence with the USCO is relevant only if there are underlying issues with the issuance of a copyright. *See Brandir Int'l, Inc. v. Cascade Pac. Lumber Co.*, 834 F.2d 1142, 1146 (2d Cir. 1987) (citing the copyright office's correspondence because it *refused* registration). In *Aalmuhammed v. Lee*, 202 F.3d 1227, 1230, 1234-36 (9th Cir. 2000), cited by OpenAI, the court assessing the plaintiff's claim of co-authorship of a joint work relied on evidence of the plaintiff's relationship with the purported co-authors to rebut the presumptive validity of his copyright registration—not his USCO correspondence. Absent factual or legal predicate establishing the relevance, OpenAI has not shown discovery of USCO correspondence with The Times over the past century is warranted. Fed. R. Civ. P. 26(b)(1).

---

[3] The allegations in this case concern infringement of written content, and therefore, the far more frequent complaints news organizations receive relating to images would not be within the scope of permissible discovery. *See Usherson v. Bandshell Artist Mgmt.*, 2020 WL 3483661, at *1-2 (S.D.N.Y. June 26, 2020) (detailing Richard Liebowitz's thousands of frivolous copyright infringement complaints over photos).

July 3, 2024
Page 4

Respectfully submitted,

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

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

cc: All Counsel of Record (via ECF)