orrick

Orrick, Herrington & Sutcliffe LLP
The Orrick Building
405 Howard Street
San Francisco, CA 94105-2669

+1 415 773 5700

orrick.com

November 18, 2024

Sent via ECF

Annette L. Hurst

E ahurst@orrick.com
D +1 415 773 4585
F +1 415 773 5759

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl Street
New York, NY 10007-1312

RE: *The New York Times Co. v. Microsoft Corp., et al*. No. 1:23-cv-11195
       Daily News, LP, et al. v. Microsoft Corp., et al., No. 1:24-cv-3285
       The Center for Investigative Reporting, Inc. v. Microsoft Corp., et al. No. 1:24-cv-04872

Dear Judge Wang:

Microsoft respectfully requests a pre-motion conference regarding Plaintiff The New York Times Company's ("The Times") refusal to produce documents concerning The Times's own use of Generative AI. Specifically, Microsoft requests the Court compel The Times to produce: (1) documents reflecting efforts by The Times to train, fine-tune, or otherwise develop and use a Generative AI tool; and (2) a survey conducted by The Times concerning workplace generative AI usage. These documents are highly relevant to three aspects of Microsoft's defense: *first*, they are relevant to establishing the fair use first factor—namely, that these models are a transformative use; *second*, they are directly relevant to assessing The Times's claims of harm under the Supreme Court's decision in *Google v. Oracle*; and *third*, they are relevant to establish Microsoft's substantial noninfringing uses defense to The Times's claims for contributory liability. Despite this relevance, The Times has nevertheless refused to produce relevant, non-privileged documents about its own efforts to create and its own use of Generative AI tools. Microsoft asks the Court to compel The Times to produce these documents.

The evidence Microsoft requests is highly relevant to three significant aspects of its defense:

1.   *Transformative Use.* Microsoft is entitled to this evidence to prove that Generative AI models are a highly transformative use of the works in suit. The process of turning a mass of text into a Generative AI tool, through training an AI using (among others) the works in suit, is clearly a significant transformation. The evidence Microsoft seeks therefore relates to the "primary inquiry" underpinning the first factor: "whether the new work, '. . . merely supersede[s] the objects of the original creation, . . . or instead adds something new, with a further purpose . . . . [I]t asks, in other words, whether and to what extent the new work is 'transformative.'" *Authors Guild v. Google, Inc.*, 804 F.3d 202, 214-15 (2d Cir. 2015) (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-79 (1994)). Additionally, with respect to The Times's output claims regarding Microsoft's generative search engines, the transformative purpose of delivering information from many sources more efficiently is also significant: "[A] secondary use may be a fair use if it utilizes

1

technology to achieve the transformative purpose of improving the efficiency of delivering content without unreasonably encroaching on the commercial entitlements of the rights holder." *Fox News Network, LLC v. TVEyes*, 883 F.3d 169, 177 (2d Cir. 2018).

2.   *Lack of Harm.* Additionally, this evidence is relevant in two ways to the fourth factor of fair use. Any failed efforts by The Times to develop its own Generative AI system using its works would undermine its claims of harm. *Google LLC v. Oracle America, Inc.*, 593 U.S. 1, 36 (2021) (failure to develop technology weighed against harm in analysis of fourth factor). Positive uses of the technology are also relevant to balancing any claims of harm under the fourth factor of fair use. *See id.* at 35 ("we must take into account the public benefits the copying will likely produce").

3.   *Substantial Noninfringing Uses.* Finally, in addition to supporting Microsoft's showing of transformative use under the first factor, the nature of the uses of the works at issue is also relevant to the substantial noninfringing uses defense. The Times alleges that Microsoft is secondarily liable in two respects: (1) that Microsoft is secondarily liable for OpenAI's development of AI; and (2) that Microsoft is secondarily liable for any use by end users of its tools that result in infringement of its works (output infringement). That the technology is capable of substantial noninfringing uses is a defense to both of these claims. *See Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 417 (1984) (holding early videocassette recorders are "capable of substantial noninfringing uses" based on evidence of "[p]rivate, noncommercial time-shifting in the home"); *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*, 536 F.3d 121, 133 (2d Cir. 2008) ("asking whether copying copyrighted material is only 'incidental' to a given technology is akin to asking whether that technology has 'commercially significant noninfringing uses,' another inquiry the *Sony* Court found relevant to whether imposing contributory liability was just.") (quoting *Sony*, 464 U.S. at 442).

The two categories of requested documents directly address these aspects of Microsoft's defenses because The Times has *itself* attempted to create Generative AI, and is also apparently making widespread use of it in the workplace.

1.   *The Times's Effort to Develop Generative AI Tools.* Based upon public reports indicating that The Times tried to train its own Generative AI model, Microsoft propounded the following requests: Request No. 31 (training and fine-tuning of The Times's "generative AI ad tool" referenced in an article published by Axios); No. 32 (training and fine-tuning of any other Generative AI model developed, used, or considered for use by The Times); No. 102 (The Times's efforts to train, develop, or use any LLM or Generative AI model); No. 106 (The Times's ██████ ████████████████████. *See* Ex. A (The Times's Resps. & Objs. to Microsoft's First Set of Requests for Production) at 23-24; Ex. B (The Times' Resps. & Objs. to Microsoft's Second Set of Requests for Production) at 20-23. The Times's efforts to develop its own models are directly relevant to the first and fourth factors. The Times has outright refused these requests.

Nevertheless, documents confirming Microsoft's theory have leaked into The Times's production. One such document indicates that The Times ███████████████████████████████ (Ex. C, NYT_00198239), and that ██:

[BLACK REDACTION BOX]

Accordingly, The Times must produce all information about its own efforts to construct AI models. Whether and how The Times developed and is using its own Generative AI tool is directly relevant to the fair use and substantial non-infringing use defenses. Microsoft's Requests are not overly broad—they are focused on the training, fine-tuning, development, or use of any such models. To the extent that this is limited to The Times "generative AI ad tool" and ▮▮▮▮▮▮▮▮▮▮ ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ this should not be overly burdensome. And, indeed, the thrust of The Times pushback on these Requests has not been about burden, but instead about relevance. This material is unquestionably relevant, and The Times should be compelled to produce it. *See, e.g., Viacom Int'l Inc. v. Youtube Inc.*, 253 F.R.D. 256, 261-62 (S.D.N.Y. 2008) (granting motion to compel evidence of substantial non-infringing uses).

2.   *Documents Regarding The Times's Workplace Use of GenAI.* [BLACK REDACTION BOX]

[BLACK REDACTION BOX] Indeed, on The Times's podcast Hard Fork, reporter Kevin Rouse described generative AI as "the best research assistant [he has] ever had." *Can A.I. Save the Middle Class*?, Hard Fork Ep. 77, available at YouTube https://www.youtube.com/watch?v=Xp1Bu-ssFVg. Documents produced by The Times indicate that ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ ▮▮▮▮▮▮▮▮▮▮ Ex. D, NYT_00115513. Microsoft therefore served Request No. 108 seeking documents and communications relating to ▮▮▮▮▮▮ Ex. B at 23-24. Understanding how The Times's employees themselves are using Generative AI tools goes directly to the question of whether these tools are capable of substantial non-infringing uses, as well as the societal benefits they offer. It is clear that The Times has in its possession information regarding ▮▮▮▮▮

Indeed, it appears that The Times was in fact ▮▮▮▮▮▮ ▮▮▮▮ that it seeks to dismantle with this lawsuit. *See* Ex. D ▮▮▮▮▮▮▮

▮▮▮▮▮▮▮▮▮▮▮ Production of this ▮▮▮▮ is the least invasive and most efficient means for Microsoft to discover the evidence to which it is entitled of how ▮▮▮▮▮ ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮.

The Times has not articulated any burden associated with the production of any of these documents—it has simply stated that it will not produce them because it believes its development and use of such tools is irrelevant—and has otherwise failed to provide any reasonable or rational explanation to justify excusing it from participating in standard fair use discovery.

For the reasons presented, Microsoft respectfully requests the Court compel The Times to produce documents responsive to Request Nos. 31-32, 102, 106, and 108.

Respectfully submitted,

/s/ Annette L. Hurst

Annette L. Hurst