

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

THE NEW YORK TIMES COMPANY

Plaintiff,

v.

MICROSOFT CORPORATION, OPENAI, INC.,
OPENAI LP, OPENAI GP, LLC, OPENAI, LLC,
OPENAI OPCO LLC, OPENAI GLOBAL LLC,
OAI CORPORATION, LLC, and OPENAI
HOLDINGS, LLC,

Defendants.

Civil Action No. _____

COMPLAINT

JURY TRIAL DEMANDED

Plaintiff The New York Times Company (“The Times”), by its attorneys Susman Godfrey LLP and Rothwell, Figg, Ernst & Manbeck, P.C., for its complaint against Defendants Microsoft Corporation (“Microsoft”) and OpenAI, Inc., OpenAI LP, OpenAI GP LLC, OpenAI LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC, OpenAI Holdings, LLC, (collectively “OpenAI” and, with Microsoft, “Defendants”), alleges as follows:

I. NATURE OF THE ACTION

1. Independent journalism is vital to our democracy. It is also increasingly rare and valuable. For more than 170 years, The Times has given the world deeply reported, expert, independent journalism. Times journalists go where the story is, often at great risk and cost, to inform the public about important and pressing issues. They bear witness to conflict and disasters, provide accountability for the use of power, and illuminate truths that would otherwise go unseen. Their essential work is made possible through the efforts of a large and expensive organization that provides legal, security, and operational support, as well as editors who ensure their journalism meets the highest standards of accuracy and fairness. This work has always been important. But

within a damaged information ecosystem that is awash in unreliable content, The Times’s journalism provides a service that has grown even more valuable to the public by supplying trustworthy information, news analysis, and commentary.

2. Defendants’ unlawful use of The Times’s work to create artificial intelligence products that compete with it threatens The Times’s ability to provide that service. Defendants’ generative artificial intelligence (“GenAI”) tools rely on large-language models (“LLMs”) that were built by copying and using *millions* of The Times’s copyrighted news articles, in-depth investigations, opinion pieces, reviews, how-to guides, and more. While Defendants engaged in widescale copying from many sources, they gave Times content particular emphasis when building their LLMs—revealing a preference that recognizes the value of those works. Through Microsoft’s Bing Chat (recently rebranded as “Copilot”) and OpenAI’s ChatGPT, Defendants seek to free-ride on The Times’s massive investment in its journalism by using it to build substitutive products without permission or payment.

3. The Constitution and the Copyright Act recognize the critical importance of giving creators exclusive rights over their works. Since our nation’s founding, strong copyright protection has empowered those who gather and report news to secure the fruits of their labor and investment. Copyright law protects The Times’s expressive, original journalism, including, but not limited to, its millions of articles that have registered copyrights.

4. Defendants have refused to recognize this protection. Powered by LLMs containing copies of Times content, Defendants’ GenAI tools can generate output that recites Times content verbatim, closely summarizes it, and mimics its expressive style, as demonstrated by scores of examples. *See* Exhibit J. These tools also wrongly attribute false information to The Times.

5. Defendants also use Microsoft’s Bing search index, which copies and categorizes The Times’s online content, to generate responses that contain verbatim excerpts and detailed summaries of Times articles that are significantly longer and more detailed than those returned by traditional search engines. By providing Times content without The Times’s permission or authorization, Defendants’ tools undermine and damage The Times’s relationship with its readers and deprive The Times of subscription, licensing, advertising, and affiliate revenue.

6. Using the valuable intellectual property of others in these ways without paying for it has been extremely lucrative for Defendants. Microsoft’s deployment of Times-trained LLMs throughout its product line helped boost its market capitalization by a trillion dollars in the past year alone. And OpenAI’s release of ChatGPT has driven its valuation to as high as \$90 billion. Defendants’ GenAI business interests are deeply intertwined, with Microsoft recently highlighting that its use of OpenAI’s “best-in-class frontier models” has generated customers—including “leading AI startups”—for Microsoft’s Azure AI product.¹

7. The Times objected after it discovered that Defendants were using Times content without permission to develop their models and tools. For months, The Times has attempted to reach a negotiated agreement with Defendants, in accordance with its history of working productively with large technology platforms to permit the use of its content in new digital products (including the news products developed by Google, Meta, and Apple). The Times’s goal during these negotiations was to ensure it received fair value for the use of its content, facilitate the continuation of a healthy news ecosystem, and help develop GenAI technology in a responsible way that benefits society and supports a well-informed public.

¹ *Microsoft Fiscal Year 2024 First Quarter Earnings Conference Call*, MICROSOFT INVESTOR RELATIONS (Oct. 24, 2023), <https://www.microsoft.com/en-us/Investor/events/FY-2024/earnings-fy-2024-q1.aspx>.

8. These negotiations have not led to a resolution. Publicly, Defendants insist that their conduct is protected as “fair use” because their unlicensed use of copyrighted content to train GenAI models serves a new “transformative” purpose. But there is nothing “transformative” about using The Times’s content without payment to create products that substitute for The Times and steal audiences away from it. Because the outputs of Defendants’ GenAI models compete with and closely mimic the inputs used to train them, copying Times works for that purpose is not fair use.

9. The law does not permit the kind of systematic and competitive infringement that Defendants have committed. This action seeks to hold them responsible for the billions of dollars in statutory and actual damages that they owe for the unlawful copying and use of The Times’s uniquely valuable works.

II. JURISDICTION AND VENUE

10. The Court has subject matter jurisdiction under 28 U.S.C. §§ 1331 and 1338(a) because this action arises under the Copyright Act of 1976, 17 U.S.C. § 101, et seq.

11. Jurisdiction over Microsoft and OpenAI is proper because they have purposely availed themselves of the privilege of conducting business in New York. A substantial portion of Microsoft and OpenAI’s widespread infringement and other unlawful conduct alleged herein occurred in New York, including the distribution and sales of Microsoft and OpenAI’s Generative Pre-training Transformer (“GPT”)-based products like ChatGPT, ChatGPT Enterprise, Bing Chat, Azure OpenAI Service, Microsoft 365 Copilot, and related application programming interface (API) tools within New York to New York residents. Furthermore, both Microsoft and the OpenAI Defendants maintain offices and employ personnel in New York who, upon information and belief, were involved in the creation, maintenance, or monetization of Microsoft and OpenAI’s widespread infringement and other unlawful conduct alleged herein.

12. Because The Times’s principal place of business and headquarters is in this District, the injuries alleged herein from Microsoft and OpenAI’s widespread infringement and other unlawful conduct foreseeably occurred in this District.

13. Venue is proper under 28 U.S.C. § 1400(a) because Defendants or their agents reside or may be found in this District, through the infringing and unlawful activities—as well as Defendants’ sales and monetization of such activity—that occurred in this District. Venue is also proper under 28 U.S.C. § 1391(b)(2) because a substantial part of the events giving rise to The Times’s claims occurred in this District, including the marketing, sales, and licensing of Defendants’ GenAI products built on the infringement of The Times’s intellectual property within this District. Upon information and belief, OpenAI has sold subscriptions for ChatGPT Plus to New York residents, and both Microsoft and OpenAI enjoy a substantial base of monthly active users of Bing Chat and ChatGPT in New York. OpenAI has licensed its GPT models to New York residents and companies headquartered in New York. For example, this year, OpenAI struck deals to license its GPT models to the Associated Press (AP) and Morgan Stanley, both companies headquartered in New York.

III. THE PARTIES

14. Plaintiff The New York Times Company is a New York corporation with its headquarters and principal place of business in New York. The Times publishes digital and print products, including its core news product, The New York Times, which is available on its mobile applications, on its website (NYTimes.com), and as a printed newspaper, and associated content such as its podcasts. The Times also publishes other interest-specific publications, including The Athletic (sports media), Cooking (recipes and other cooking-related content), Games (puzzles and

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.