

The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work

Millions of articles from The New York Times were used to train chatbots that now compete with it, the lawsuit said.



By Michael M. Grynbaum and Ryan Mac

Dec. 27, 2023

The New York Times sued OpenAI and Microsoft for copyright infringement on Wednesday, opening a new front in the increasingly intense legal battle over the unauthorized use of published work to train artificial intelligence technologies.

The Times is the first major American media organization to sue the companies, the creators of ChatGPT and other popular A.I. platforms, over copyright issues associated with its written works. The lawsuit, filed in Federal District Court in Manhattan, contends that millions of articles published by The Times were used to train automated chatbots that now compete with the news outlet as a source of reliable information.

The suit does not include an exact monetary demand. But it says the defendants should be held responsible for “billions of dollars in statutory and actual damages” related to the “unlawful copying and use of The Times’s uniquely valuable works.” It also calls for the companies to destroy any chatbot models and training data that use copyrighted material from The Times.

In its complaint, The Times said it approached Microsoft and OpenAI in April to raise concerns about the use of its intellectual property and explore “an amicable resolution,” possibly involving a commercial agreement and “technological guardrails” around generative A.I. products. But it said the talks had not produced a resolution.

An OpenAI spokeswoman, Lindsey Held, said in a statement that the company had been “moving forward constructively” in conversations with The Times and that it was “surprised and disappointed” by the lawsuit.

“We respect the rights of content creators and owners and are committed to working with them to ensure they benefit from A.I. technology and new revenue models,” Ms. Held said. “We’re hopeful that we will find a mutually beneficial way to work together, as we are doing with many other publishers.”

Microsoft declined to comment on the case.

The lawsuit could test the emerging legal contours of generative A.I. technologies — so called for the text, images and other content they can create after learning from large data sets — and could carry major implications for the news industry. The Times is among a small number of outlets that have built successful business models from online journalism, but dozens of newspapers and magazines have been hobbled by readers’ migration to the internet.

At the same time, OpenAI and other A.I. tech firms — which use a wide variety of online texts, from newspaper articles to poems to screenplays, to train chatbots — are attracting billions of dollars in funding.

OpenAI is now valued by investors at more than \$80 billion. Microsoft has committed \$13 billion to OpenAI and has incorporated the company’s technology into its Bing search engine.

“Defendants seek to free-ride on The Times’s massive investment in its journalism,” the complaint says, accusing OpenAI and Microsoft of “using The Times’s content without payment to create products that substitute for The Times and steal audiences away from it.”

The defendants have not had an opportunity to respond in court.

Concerns about the uncompensated use of intellectual property by A.I. systems have coursed through creative industries, given the technology's ability to mimic natural language and generate sophisticated written responses to virtually any prompt.

The actress Sarah Silverman joined a pair of lawsuits in July that accused Meta and OpenAI of having “ingested” her memoir as a training text for A.I. programs. Novelists expressed alarm when it was revealed that A.I. systems had absorbed tens of thousands of books, leading to a lawsuit by authors including Jonathan Franzen and John Grisham. Getty Images, the photography syndicate, sued one A.I. company that generates images based on written prompts, saying the platform relies on unauthorized use of Getty's copyrighted visual materials.

The boundaries of copyright law often get new scrutiny at moments of technological change — like the advent of broadcast radio or digital file-sharing programs like Napster — and the use of artificial intelligence is emerging as the latest frontier.

Our business reporters. Times journalists are not allowed to have any direct financial stake in companies they cover.

“A Supreme Court decision is essentially inevitable,” Richard Tofel, a former president of the nonprofit newsroom ProPublica and a consultant to the news business, said of the latest flurry of lawsuits. “Some of the publishers will settle for some period of time — including still possibly The Times — but enough publishers won't that this novel and crucial issue of copyright law will need to be resolved.”

Microsoft has previously acknowledged potential copyright concerns over its A.I. products. In September, the company announced that if customers using its A.I. tools were hit with copyright complaints, it would indemnify them and cover the associated legal costs.

Other voices in the technology industry have been more steadfast in their approach to copyright. In October, Andreessen Horowitz, a venture capital firm and early backer of OpenAI, wrote in comments to the U.S. Copyright Office that exposing A.I. companies to copyright liability would “either kill or significantly hamper their development.”

“The result will be far less competition, far less innovation and very likely the loss of the United States' position as the leader in global A.I. development,” the investment firm said in its statement.

Besides seeking to protect intellectual property, the lawsuit by The Times casts ChatGPT and other A.I. systems as potential competitors in the news business. When chatbots are asked about current events or other newsworthy topics, they can generate answers that rely on journalism by The Times. The newspaper expresses concern that readers will be satisfied with a response from a chatbot and decline to visit The Times's website, thus reducing web traffic that can be translated into advertising and subscription revenue.

The complaint cites several examples when a chatbot provided users with near-verbatim excerpts from Times articles that would otherwise require a paid subscription to view. It asserts that OpenAI and Microsoft placed particular emphasis on the use of Times journalism in training their A.I. programs because of the perceived reliability and accuracy of the material.

Media organizations have spent the past year examining the legal, financial and journalistic implications of the boom in generative A.I. Some news outlets have already reached agreements for the use of their journalism: The Associated Press struck a licensing deal in July with OpenAI, and Axel Springer, the German publisher that owns Politico and Business Insider, did likewise this month. Terms for those agreements were not disclosed.

The Times is exploring how to use the nascent technology itself. The newspaper recently hired an editorial director of artificial intelligence initiatives to establish protocols for the newsroom's use of A.I. and examine ways to integrate the technology into the company's journalism.

In one example of how A.I. systems use The Times's material, the suit showed that Browse With Bing, a Microsoft search feature powered by ChatGPT, reproduced almost verbatim results from Wirecutter, The Times's product review site. The text results from Bing, however, did not link to the Wirecutter article, and they stripped away the referral links in the text that Wirecutter uses to generate commissions from sales based on its recommendations.

"Decreased traffic to Wirecutter articles and, in turn, decreased traffic to affiliate links subsequently lead to a loss of revenue for Wirecutter," the complaint states.

The lawsuit also highlights the potential damage to The Times's brand through so-called A.I. "hallucinations," a phenomenon in which chatbots insert false information that is then wrongly attributed to a source. The complaint cites several cases in which Microsoft's Bing Chat provided incorrect information that was said to have come from The Times, including results for "the 15 most heart-healthy foods," 12 of which were not mentioned in an article by the paper.

"If The Times and other news organizations cannot produce and protect their independent journalism, there will be a vacuum that no computer or artificial intelligence can fill," the complaint reads. It adds, "Less journalism will be produced, and the cost to society will be enormous."

The Times has retained the law firms Susman Godfrey and Rothwell, Figg, Ernst & Manbeck as outside counsel for the litigation. Susman represented Dominion Voting Systems in its defamation case against Fox News, which resulted in a \$787.5 million settlement in April. Susman also filed a proposed class action suit last month against Microsoft and OpenAI on behalf of nonfiction authors whose books and other copyrighted material were used to train the companies' chatbots.

Benjamin Mullin contributed reporting.

Michael M. Grynbaum writes about the intersection of media, politics and culture. He has been a media correspondent at The Times since 2016. More about Michael M. Grynbaum

Ryan Mac covers corporate accountability across the global technology industry. More about Ryan Mac

A version of this article appears in print on , Section B, Page 1 of the New York edition with the headline: The Times Sues OpenAI and Microsoft