

## Big Tech Goes on Trial, Part 3: Can OpenAI Survive Copyright Lawsuits?

Santiago Garcia-Mata

In late 2023, The New York Times filed suit against OpenAI and Microsoft for copyright infringement. New York Times is alleging that OpenAI and Microsoft are stealing its copyrighted content to train their large language models and profiting from it. The New York Times hopes to establish the wrongfulness of this conduct. Depending on who you ask, the stakes of the lawsuit are very high, with the potential to change the future of the news business, copyright law, innovation, and of OpenAI and other generative AI companies.

The New York Times' case focuses on OpenAI's flagship AI language model, ChatGPT. When a user asks ChatGPT a question, there is a good chance that some of the information in the AI's response will be sourced from a New York Times article. The reason for this is that the Large Language Model that powers ChatGPT has been trained on over 500 gigabytes of data, including newspaper archives containing works of The New York Times. This training powers generative AI, allowing it to effectively respond to prompts and for the companies behind it to reap massive profits.

Specifically, The New York Times alleges that OpenAI trained its model with copyrighted Times content and did not pay proper licensing fees. Thus, they claim that OpenAI is allowed to compete with and closely mimic the New York Times by summarizing its work or copying recommendations from its Wirecutter reviews. Additionally, the lawsuit alleges OpenAI regurgitates Times articles verbatim, providing over 100 examples of such occurrences when using ChatGPT. According to one estimate, the New York Times could be looking to recover \$450 billion in damages due to the millions of implicated articles and the cost per instance of copying.

On the other side, OpenAI and Microsoft deny the allegations, claiming that The New York Times intentionally manipulated prompts to provide examples of regurgitation. OpenAI claims that regurgitation is a rare bug that the company is working to eliminate. The company has also proven to be willing to pay licensing fees for the model's training, even negotiating deals with the Associated Press and Axel Springer. OpenAI was even negotiating a licensing deal with The New York Times prior to the lawsuit, though it was never completed. Additionally, OpenAI is hoping to draw parallels to the Supreme Court victory of Sony against Universal Studios in 1984 concerning VCR recording, arguing that technological innovation should not be overly hampered by copyright owners.

The lawsuit against OpenAI also resembles a famous copyright suit against Napster. The peer-to-peer file-sharing site made downloading music easy and free, provoking record companies to sue for copyright infringement. In 2001, a federal court ruled Napster liable for copyright infringement, marking a big win against a large technology company. Though OpenAI's legal battle is different than Napster's, the precedent set by this case will have similarly important implications for copyright law.

Technological innovation has long been followed by people wanting to protect their creations. When technology made copying easier over 300 years ago, the earliest copyright laws also emerged as a way for authors to protect their intellectual property. Though artificial intelligence is a large technological leap, and one that is still in its earliest phases, its future will be impacted by this development in copyright law.

For more information, see the full article [here](#).