

AI LAWSUIT MAY ECLIPSE CLAIMS OF FAIR USE

Hodgson Russ Intellectual Property Litigation Alert
April 5, 2024

Does an artificial intelligence (“AI”) company have the right to use copyrighted content to train its AI model? That question is being reviewed where two giants have squared off in a case where the New York Times sued OpenAI and Microsoft, the companies behind ChatGPT, for copyright infringement in federal court in New York City.

The New York Times registers the copyright in its print edition every day, it maintains a paywall for certain content, and its terms of service limits the ability to copy its content. The New York Times has licensed its content for commercial uses in the past, but the New York Times and OpenAI do not have a licensing agreement.

Some technical background sets the stage for the claims by the New York Times. ChatGPT is a large language AI model (“LLM”). LLMs use probabilities to predict which word will follow the set of preceding words, much like how weather models use past patterns to predict tomorrow’s weather. An LLM needs large datasets to generate accurate probabilities. OpenAI uses datasets that contain a copy of all the text published online, which includes content created by the New York Times.

The New York Times alleges that industrial-scale copyright infringement fueled ChatGPT’s success. Content from the New York Times appears frequently in the datasets used to train ChatGPT. Additionally, with creative prompting, ChatGPT fully reproduces paywalled content. According to the New York Times, the unlicensed use of New York Times content in training ChatGPT and the ability to reproduce paywalled content, deprives the New York Times of the revenue necessary to produce high quality journalism.

Microsoft and OpenAI moved to dismiss some of the claims made by the New York Times. In its motion, Microsoft alleges that the fair use doctrine protects their conduct. The fair use doctrine allows certain unlicensed reproduction of copyrighted works for beneficial activities like education and scientific advancement. Microsoft compares LLMs to other technologies that allow users to lawfully reproduce copyrighted content, such as VCRs, musical instruments, and search engines. According to Microsoft, content from the New York Times is merely an input to train ChatGPT; ChatGPT does not supplant the market for content from the New York Times.

Attorneys

Alfonzo Cutaia
John DiMaio Ph.D.
Emily Florczak
Rob Fluskey
Neil Friedman
Jodyann Galvin
Joshua Levitt
John Lopinski Ph.D.
Nathaniel Lucek
Fallon Martin
Christopher Massaroni
Ryan McGonigle
Scott Paton
Charles Rauch
R. Kent Roberts
Paul Roman Jr. Ph.D.
Alexandria Rowen
Gary Schober
George Snyder
Melissa Subjeck
James Zawodzinski Jr.

Practices & Industries

Intellectual Property Litigation

AI LAWSUIT MAY ECLIPSE CLAIMS OF FAIR USE

In its response to the motions to dismiss, the New York Times argues that ChatGPT copies and then transforms news articles to present a market alternative to traditional journalism and do not simply reproduce content like a VCR or search engine. The New York Times urges the Court to read copyright law in a way that allows both journalism and artificial intelligence to flourish.

Notably, on April 1, 2024, the judge presiding over the case refused to permit a group of authors and journalists with pending California cases against Microsoft and Open AI to intervene in the New York Times case.

This case may re-define the boundaries of fair use. The New York Times litigation has been consolidated with other New York City-based copyright lawsuits, including one brought by the Authors Guild. If Microsoft and OpenAI prevail, AI companies could use copyrighted content to train their models without a license, and creative AI users could bypass paywalls. If the New York Times prevails, AI companies may have to negotiate licenses for content to train an LLM.

Disclaimer:

This client alert is a form of attorney advertising. Hodgson Russ LLP provides this information as a service to its clients and other readers for educational purposes only. Nothing in this client alert should be construed as, or relied upon, as legal advice or as creating a lawyer-client relationship.