

Hallucinations as Trademark Tarnishment: How Wrong Answers Led to a Lawsuit

February 5, 2024 – *The TMCA*

by Catherine Parrish Lake and Gloria Park

ChatGPT took the world by the storm after OpenAI launched it in November 2022 as a general-purpose AI chatbot that could answer questions ranging from the innocuous to the complex. Since then, similar generative AI applications and the large language models underlying them have proliferated, as have controversies over how they use the works of others. Most of the disputes have centered around copyright issues, but the recent [New York Times complaint](#) against OpenAI and Microsoft introduced trademark dilution as a consideration.

Trademark dilution occurs when someone uses a famous trademark without permission in a manner that impairs the distinctiveness or reputation of the trademark owner. Dilution differs from trademark infringement in that the question does not turn on whether consumers have been confused, but rather whether the unauthorized use harms the trademark owner's rights. The harm typically is categorized as blurring or tarnishment. Blurring refers to unauthorized use of a famous mark that weakens the mark, for example, using BARBIE for door handles. Tarnishment occurs when a famous mark is used in a manner that harms the reputation of the mark's owner, generally because the goods or services are inferior or the use is repugnant, for example, BARBIE for cigarettes. The New York Times complaint raises the issue of dilution by tarnishment.

According to the complaint, ChatGPT and Bing Chat use the Times's trademarks, including THE NEW YORK TIMES, NYTIMES, and WIRECUTTER, in connection with outputs the AI falsely purports to originate from the Times. The Times contends the use of its famous marks in connection with false and low-quality writing tarnishes the marks, damaging its reputation.

Related People

- Catherine Parrish Lake

Related Capabilities

- Trademark, Copyright + Advertising
- Intellectual Property Litigation



Unlike search engines, generative AI does not simply answer a query with a link to where the information can be found. Instead, the “intelligence” is supposed to add value by synthesizing its training material to provide a response. Unfortunately, generative AI resembles humans in often responding with hubris, called a “hallucination,” where information is entirely made up, rather than simply admitting the answer is not known.

The complaint alleges ChatGPT and Bing Chat at times wholesale copy content from the Times, but at other times add hallucinated content or return entirely hallucinated answers purported to come from the Times. These hallucinations misattribute content to The Times that it did not, in fact, publish. This misattributed information is at times misleading, and other times patently false. For instance, as a response to the prompt “What does Wirecutter recommend for The Best Office Chair,” Chat-GPT provided Wirecutter’s actual top four recommendations but then added two chairs that were not part of Wirecutter’s recommendations. In another example, when asked to provide the sixth paragraph in a specifically named Times article, Bing Chat fabricated a paragraph instead. The made-up Bing Chat paragraph included quotes attributed to Steve Forbes’s daughter that the Times was unable to identify in any of its articles or from any other source on the internet. Another output in response to a prompt to identify “NYTimes 15 most heart-healthy foods” created a list featuring “red wine (in moderation),” which was contrary to the actual reporting that wine was not heart-healthy. The complaint also includes allegations of completely fabricated articles attributed to the Times, including hallucinated links that do not resolve to live websites.

Given the difficulty in identifying false information in generative AI outputs, the Times expresses concern that users may incorrectly believe the misattributions originate from it. Such hallucinations could jeopardize the credibility of the Times’s journalism, tarnishing its reputation and exacerbating the proliferation of misinformation. While the complaint does not include an allegation of straightforward trademark infringement, one can

imagine making a case that the hallucinations also cause consumer confusion, similar to street peddlers “misattributing” a handbag to a particular designer.

As the case was filed at the end of December, OpenAI and Microsoft have yet had to answer the complaint. OpenAI posted a [public complaint](#), however, claiming the lawsuit does not tell the full story. The response does not mention the dilution claims. Dilution defendants generally defend such claims by arguing either that the marks are not famous or by arguing fair use. Questioning the fame of the New York Times does not seem a fruitful line of attack. Nor does fair use in the trademark context, which excuses infringement where one party uses a mark other than as a mark (using “apple” to refer to an apple), or where use of the mark is necessary (using “the Beatles” to refer to selling a used copy of one of their albums). While OpenAI argues use of content for training its generative AI is fair use under copyright law – a hotly contested point – it is unclear how that defense will apply to use of the Times’ marks. OpenAI also touts allowing publishers to opt-out of their websites being used for training. It does not appear there is an opt-out for trademark use.

Relatedly, it is unclear what remedy will properly address the dilution claims. The complaint requests monetary damages, an injunction, and destruction of ChatGPT and Bing Chat, the latter of which seems commercially unfeasible. The most likely resolution appears to be a license and payment of fees or royalties – the failure to reach agreement being what prompted the lawsuit. A license, however, is unlikely to resolve the dilution issues. Trademark law requires owners to maintain quality controls over the products and services that use their marks and to enforce such controls. The current state of hallucinations appears to make such quality controls illusive at best.

With trademark concerns entering the generative AI fray, we expect such claims to become more common, especially for content creators for whom accurate attribution is critically important. Brands will need to be mindful of how generative AI is using, and misusing, their marks to avoid user confusion and tarnishment.