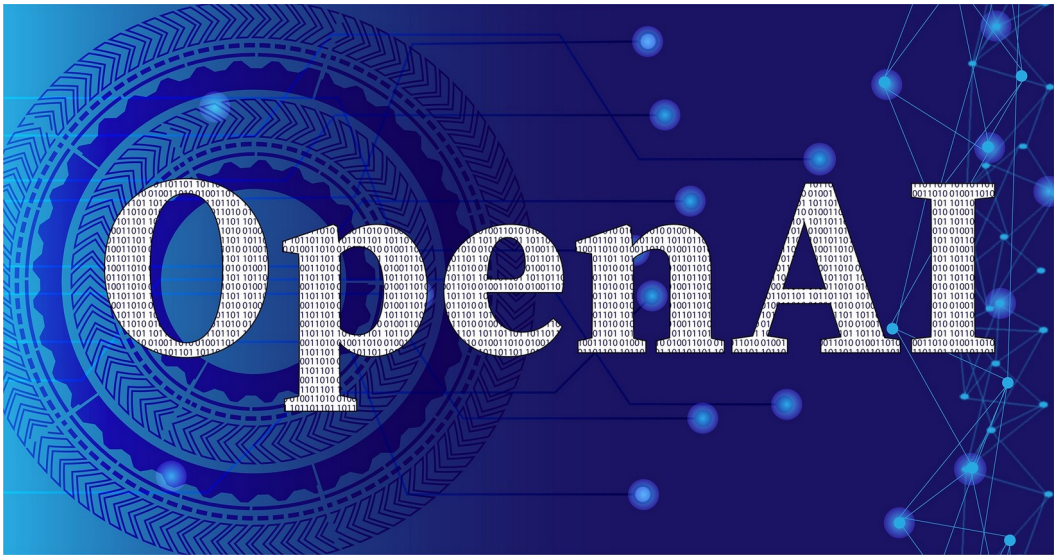
DORSEY
DORSEY+WHITNEY LLP

# Uncle Sam Weighs In: the DOJ Says Training AI on Your NYT Subscription Is Totally Fine

*November 17, 2026 – The TMCA*
by Kaleb McNeely




## Related People

- Kaleb McNeely

While various titans of AI industry have recently sounded the alarm and suggested AI development should slow down, the U.S. federal government continues to champion AI. Now the U.S. Department of Justice has waded into a civil litigation about whether OpenAI can hoover up the entire internet (including every *New York Times* article ever written) to train ChatGPT without paying a licensing fee. On September 1, 2026, the DOJ filed a Statement of Interest in the sprawling multidistrict litigation captioned *In re: OpenAI, Inc. Copyright Infringement Litigation*, essentially telling a federal judge in Manhattan: Hey, we've got some thoughts about this whole AI-training-as-copyright-infringement thing. Spoiler alert—we think it's fine.

Here's the background: The *New York Times* (and a bunch of book authors and publishers) sued OpenAI, claiming that training large language models ("LLMs") on their copyrighted works—without permission or payment—violates the Copyright Act. The plaintiffs argue that OpenAI made unauthorized copies of millions of articles, fed them into the algorithmic meat grinder, and emerged with a product that can now compete with

those very same works in the marketplace.

OpenAI's response? "We're not *republishing* your stuff; we're teaching a computer to understand language. Totally different. Also, fair use." And now the United States has weighed in—not as a party to the case, but under 28 U.S.C. § 517, a statute that lets the DOJ file statements of interest in any federal case in which it has an "interest."

The DOJ's statement opens with a not-so-subtle reminder that artificial intelligence is now a matter of national security. The DOJ warns that AI helps analyze intelligence, enhance weapons systems, and provide battlefield recommendations—and if the U.S. hamstrings its own AI industry with overly aggressive copyright enforcement, we're basically handing the keys to our adversaries. The government also points out that requiring licensing for AI training would create an oligopoly—only the biggest tech companies could afford the fees, and legacy media outlets like the *New York Times* would rake in subsidies while indie publishers get shut out. Meanwhile, the people who *use* LLMs to create art, write stories, or analyze data would be stuck in a world where only Big Tech and Big Media get to play.

Now for the legal part: the actual copyright analysis. The Copyright Act grants authors exclusive rights to reproduce and distribute their works. But it also includes a critical exception called "fair use," which allows certain unauthorized uses of copyrighted material without permission. Fair use is designed to prevent copyright from stifling the very creativity which that law is designed to foster. Section 107 of the Copyright Act lists four factors courts must consider when determining whether a use is fair: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount used; and (4) the effect on the market for the original work. No single factor is dispositive—courts must weigh them all together.

The DOJ's position? Training AI models is exceedingly transformative—so transformative, in fact, that it's a slam-dunk fair use. The first fair-use factor asks whether the secondary use adds something new, with a further purpose or different character, or whether it merely supersedes the objects of the original creation. The more transformative the use, the more likely it is fair.

The government argues that copying *New York Times* articles to train an LLM is transformative— "spectacularly so." Why? Because the purpose of the copying isn't to distribute the articles to readers—it's to teach a machine to recognize statistical patterns in language. The LLM uses the text not to duplicate the work's expressive content, but to learn vocabulary, syntax, and how words relate to each other. The output—an AI that can translate languages, edit emails, write skits, or answer complex questions—is, in the DOJ's opinion, "undoubtedly highly transformative." The DOJ analogizes to *Google LLC v. Oracle America, Inc.*, 593 U.S. 1, 35 (2021), where the Supreme Court held that Google's copying of Java code to create Android was fair use because it allowed programmers to use their existing skills in a new platform. If *that* was transformative, the government argues, then using copyrighted text to build an entirely new type of intelligent system is even more so.

The DOJ also argues that the fourth factor—the effect on the market for the original work—weighs in favor of a fair use finding. The DOJ asserts that training an LLM doesn't substitute for *New York Times* articles in any meaningful sense. Training is an internal process that doesn't "reveal" copyrighted content to the public—it just creates a tool that can *understand* language. Even if some LLM outputs *happen* to compete with journalism (like when someone asks ChatGPT to summarize the news), those outputs aren't "substantially similar" to the original articles unless they actually reproduce protected expression.

To drive the point home, the DOJ offers a hypothetical: Joan Didion once said she learned to write by typing out Ernest Hemingway's stories to see how the sentences worked. The DOJ contends that, under plaintiffs' logic, Didion would have owed Hemingway royalties every time she published an article, because her training process and her creative output were "all one use." But as the government dryly observes, to make anyone pay specifically for the use of a book each time they later draw upon it when writing new things in new ways would be unthinkable.

The DOJ's message is clear: training AI models on copyrighted text is fair use, full stop. It's transformative, it doesn't substitute for the original works, and blocking it would undermine the very purpose of copyright law—to promote "the Progress of Science and useful Arts." U.S. CONST. art. I, § 8, cl. 8.

Whether the court agrees remains to be seen. But one thing is certain: this case is about far more than whether OpenAI owes the *New York Times* a check. It's about the future of creativity, innovation, and how American law will approach the next wave of technological progress. And if nothing else, it's a reminder that even in the age of AI, the messiest legal battles still come down to a 50-year-old statute and a bunch of lawyers arguing about what "transformative" means.