UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

IN RE:

OPENAI, INC.
COPYRIGHT INFRINGEMENT LITIGATION

25-md-3143 (SHS) (OTW)

Brief of Ron Suskind and Brian L. Frye
As Amicus Curiae in Support of Neither Party

Brian L. Frye (BF1028)
838 Pacific Avenue
New Orleans, LA 70114
(917) 273-2382
brianlfrye@gmail.com

Rule 29 Statement

No party or party's counsel authored this brief in whole or in part or contributed money that was intended to fund preparing or submitting the brief. No person other than amici and their counsel contributed money that was intended to fund preparing or submitting the brief. See Fed. R. App. P. 29(a)(4)(E).

Statement of Identity and Interests of Amicus Curiae

Ron Suskind is a Pulitzer Prize-winning journalist, author, and filmmaker. Suskind is also the CEO of Sourcebase, an AI company founded and led by journalists. Sourcebase has developed an AI model customized for media that responds to prompts with answers grounded in massive collections of source materials along with web searches of verified information, creating what might be called a verification-augmented generation. This reduces LLM hallucination by tying AI generations to vetted source materials.

Mr. Suskind lives in a middle ground between the parties. He has been a contributor of signature feature stories for the New York Times since he left his position as the Wall Street Journal's Senior National Affairs reporter in 2001. Sourcebase, built around veracity practices used by investigative reporters, is currently used by several major media organizations to connect dots and make sense of huge troves of data and documents. The platform's model and citation audit trail is attractive as well to AI for enriching and grounding the generations of LLMs, which power the source-based architecture, and the company recently received a minority investment amounting to three percent ownership from OpenAI. OAI is a passive investor and has not seen or influenced this brief in any way, nor has the New York Times.

Brian L. Frye is a Professor of Law at Tulane University School of Law. His scholarship focuses primarily on copyright law, including in relation to AI and fair use. As a copyright scholar, Professor Frye studies how the purpose of United States copyright law affects the fair use doctrine. He is interested in ensuring that courts determining whether training an AI model on copyrighted works is a fair use understand the relevant market for those works. He believes that he can help inform this court about the relevant market for journalism. He has no connection to any of the parties to this action, and was not compensated in any way for his role in writing this amicus brief.

Preliminary Statement

The News Plaintiffs ask this court to hold that training a generative AI model on copyrighted newspaper articles is not a fair use, because it harms the market for and value of those articles by facilitating the creation of market substitutes, avoiding the licensing market, and reducing the value of newspaper articles.

This amicus brief observes that generative AI does not necessarily harm the market for or value of news, and may in fact benefit the News Plaintiffs by reducing their costs and increasing the value of their brands. Newspapers are in the business of selling reliable information about the world. It is slow, expensive, and difficult to gather that information. But it is also slow, expensive, and difficult to produce newspaper articles. Generative AI can benefit newspapers by making it faster, cheaper, and easier for

them to produce newspaper articles, thereby enabling them to focus their resources on gathering information.

Newspapers are primarily in the business of selling their brand. While newspapers typically own a copyright in the articles they publish, they rarely sell those articles individually. Rather, newspapers sell access to the news, especially breaking news – local, regional, national, international – and timely analysis. News is priceless when it is fresh, but soon becomes stale and less valuable. Generative AI can help newspapers build brand value by facilitating the translation of new information into new articles and enabling newspapers to devote more resources to generating valuable new information.

Journalists are also primarily in the business of selling their brand. While journalists write articles, they rarely own the copyright in their articles, which typically belongs to the publisher. As a consequence, journalists usually do not and cannot benefit from the copyright in the articles they write, but they do benefit from the attribution of those articles. When journalists produce breaking news or timely analysis, it increases the value of their brand. Generative AI can help journalists build their brand by facilitating the production of new articles, enabling journalists to focus on reporting and analysis.

There are three groups with an interest in the outcome of this case: the News Plaintiffs, the defendants, and the public. The News Plaintiffs want to profit from their copyrighted newspaper articles, the defendants want to train their generative AI models, and the public wants access to those models. This court can ensure that the outcome of this case serves the interest of all three parties by requiring generative AI models to attribute their outputs to the newspaper articles on which the models were trained, generating profits for the News Plaintiffs and ensuring public access to generative AI.

Argument

Under United States law, the primary purpose of copyright is to benefit the public by encouraging authors to create and distribute original works of authorship the public wants to consume. The Intellectual Property Clause of the Constitution gives Congress the power "[t]o promote the progress of science . . ., by securing for limited times to authors . . . the exclusive right to their . . . writings." U.S. Const. Art. I, Sec. 8, Cl. 8. And the Supreme Court has always explicitly held that copyright is intended to benefit the public, not to benefit authors. *See*, *e.g.*, *Google LLC v. Oracle America, Inc.*, 593 U.S. 1 (2021) (holding that copyright "grants an author an exclusive right to produce his work (sometimes for a hundred years or more), not as a special reward, but in order to encourage the production of works that others might reproduce more cheaply").[1]

[1] See also *Fox Film Corp. v. Doyal*, 286 U.S. 123, 127 (1932) ("The sole interest of the United States and the primary object in conferring the monopoly lie in the general benefits derived by the public from the labors of authors."); *United States v. Paramount Pictures, Inc.*, 334 U.S. 131, 158 (1948) ("The copyright law, like the patent statutes, makes reward to the owner a secondary consideration."); *Mazer v. Stein*, 347 U.S. 201, 219 (1954) ("The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'"); *Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975) ("The immediate effect of our copyright law is to secure a fair return for an 'author's' creative labor. But the ultimate aim is, by this incentive, to stimulate artistic creativity for the general public good."); *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984) ("It is intended to motivate the creative activity of authors and inventors by the provision of a special reward, and to allow the public access to the products of their genius after

Copyright encourages authors to create and distribute original works of authorship by giving them certain exclusive rights to use those works. *See* 17 U.S.C. §§ 102(a) (granting copyright protection to "original works of authorship fixed in any tangible medium of expression") and 106 (enumerating the exclusive rights of copyright owners). But the exclusive rights of copyright owners are not absolute.

Among other things, those rights are limited by the fair use doctrine, which provides that "the fair use of a copyrighted work . . . is not an infringement of copyright." 17 U.S.C. § 107. As the Supreme Court has observed, fair use "can focus on the legitimate need to provide incentives to produce copyrighted material while examining the extent to which yet further protection creates unrelated or illegitimate harms in other markets or to the development of other products." *Google LLC v. Oracle America, Inc.*, 593 U.S. 1 (2021).

The fair use doctrine helps ensure that copyright promotes the public interest by providing that certain socially valuable uses of copyrighted works do not require the permission of the copyright owner. "We have described the 'fair use' doctrine, originating in the courts, as an 'equitable rule of reason' that 'permits courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster.'" *Google LLC v. Oracle America, Inc.*, 593 U.S. 1 (2021) (quoting *Stewart v. Abend*, 495 U. S. 207, 236 (1990)). As the fair use statute explicitly states, using a copyrighted work without permission "for purposes such as criticism, comment, news reporting, teaching . . ., scholarship, or research, is not an infringement of copyright." 17 U.S.C. § 107. Essentially, the fair use doctrine ensures that copyright promotes innovation, rather than preventing it. *See Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023) ("The fair use doctrine 'permits courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster.'") (quoting *Stewart v. Abend*, 495 U. S. 207, 236 (1990)).

I. The Fair Use Analysis

The fair use statute requires courts to consider four factors when determining whether a particular use of a copyrighted work is a fair use:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.

---

the limited period of exclusive control has expired."); *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 349 (1991) ("The primary objective of copyright iso not to reward the labor of authors, but '[t]o promote the Progress of Science and useful Arts.'"); and *Eldred v. Ashcroft*, 537 U.S. 186, 214 (2003) ("We can demur to petitioners' description of the Copyright Clause as a grant of legislative authority empowering Congress to secure a bargain—this for that.") (internal quotation marks omitted).

17 U.S.C. § 107.

According to the Supreme Court, the first factor of the fair use analysis asks whether the use is "transformative," because it adds a "new expression, meaning, or message." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994). The second factor asks whether the copied work is factual, and therefore more susceptible to fair use. *Id.* at 586. The third factor asks whether the copier used more of the work than necessary. *Id.* at 588. And the fourth factor asks whether the copying negatively affected the market for the work. *Id.* at 590.

It is generally accepted that the first and fourth factors are the most important to the fair use analysis. *Compare Harper & Row, Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539 (1985) (stating that factor four is "undoubtedly the single most important element of fair use") *with Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994) (holding that the first factor is central to the fair use inquiry). *See also* Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions Updated*, 1978-2019, 10 NYU J. Int. Prop. & Ent. L. 1 (2020) (showing that the first and fourth factors tend to determine the outcome of the fair use analysis).

Of course, the second and third factors are still important, especially in relation to news. The second factor provides that copyright protects fictional works more strongly than factual ones, which means that newspaper articles are weakly protected. And the third factor links the fair use analysis to the quantity of copying, which is minimal in AI training, which only really involves copying the algorithmic patterns or "weights" that are hidden in every text.

A. Transformative Use

The first fair use factor instructs the court to consider whether an allegedly infringing use of a copyrighted work is "transformative," because it "adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994). *See also Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023) ("A use that has a further purpose or different character is said to be 'transformative.'"). By contrast, the fourth factor instructs the court to consider whether the use of a copyrighted work has harmed or is likely to harm the market for or value of the copyrighted work. *Campbell*, 510 U.S. at 591.

Every court that has considered whether training a generative AI model on copyrighted works is a fair use has determined that AI training is a transformative use. *See*, *e.g.*, *Kadrey v. Meta*, 788 F. Supp. 3d 1026 (N.D. Ca. 2025) and *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007 (N.D. Ca. 2025) (holding that training a generative AI model is a transformative use of a copyrighted work). In *Kadrey*, the court held that using copyrighted books to train a generative AI model "was highly transformative." 788 F. Supp. 3d at 1044. And in *Bartz*, the court held that using copyrighted works to train a generative AI model "was transformative — spectacularly so." 787 F. Supp. 3d at 1021. While the Third Circuit has held that training an AI model on copyrighted elements of a legal research tool in order to create a competing legal research tool was not a transformative use, it also explicitly acknowledged that using copyrighted works to train a generative AI model could be transformative. *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intel. Inc.*, No. 25-2153, 2026 WL 2932557, at *6 n.7 (3d Cir. Sept. 29, 2026) ("Unlike the AI models in *Bartz*

and *In re: OpenAI*, ROSS's AI platform cannot generate original expression, and the evidence here supports the opposite conclusion about transformativeness.").

Training a generative AI model on copyrighted works is transformative, because it is a use the author never even contemplated that does not depend on the particular expressive content of the works. *See Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 528 (2023) (holding that "the first factor relates to the problem of substitution—copyright's bête noire."). On the contrary, generative AI training involves the analysis of millions or billions of works in order to identify hidden patterns that exist across all of those works, which generative AI models use to produce new works. *See id.* at 531 ("In a broad sense, a use that has a distinct purpose is justified because it furthers the goal of copyright, namely, to promote the progress of science and the arts, without diminishing the incentive to create."). It follows that OpenAI's use of copyrighted works to train its generative AI models is also a transformative use.

B. Market Effects

Accordingly, this amicus brief will focus on the fourth fair use factor, which requires the court to consider whether the accused use of a copyrighted work harmed the market for or value of the work. As Supreme Court has observed, the fourth fair use factor presumes harm if and only if the defendant merely reproduced the plaintiff's copyrighted work for a commercial purpose. *See Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591 (1994) ("No 'presumption' or inference of market harm . . . is applicable to a case involving something beyond mere duplication for commercial purposes."). If the defendant used the plaintiff's work in a new way or for a new purpose, then there is no presumption of harm. *Id.* And harm is even less likely if the use was transformative, like using works to train a generative AI model. *Id.* ("But when, on the contrary, the second use is transformative, market substitution is at least less certain, and market harm may not be so readily inferred.").

The fourth fair use factor requires the court to consider not only whether the accused use harmed the market for or value of the copyrighted work, but also whether copyright should care about the harm. *See, e.g., Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 593 (1994) (holding that "the only harm to derivatives that need concern us, as discussed above, is the harm of market substitution"). While many copyright owners want to control every conceivable use of their works, the fair use doctrine says they are only entitled to control certain uses, with additional limitations.

The News Plaintiffs claim that OpenAI's use of their newspaper articles to train its generative AI models harmed the market for and value of those articles by facilitating the creation of newspaper articles, discouraging the licensing of newspaper articles, and diluting the market for newspaper articles.

All of those claims are superficially true. OpenAI's use of newspaper articles to train its generative AI models did harm the market for and value of those articles. It did facilitate the creation of newspaper articles by helping journalists write them. It did discourage licensing newspaper articles for AI training by showing it isn't required. And it did dilute the market for newspaper articles by creating a lot more of them.

But none of those harms are relevant, because they were all the result of fair competition, which copyright is supposed to encourage, not prohibit. We want journalists to write newspaper articles. We want people to use newspaper articles in transformative ways. And we want newspapers to publish more articles. While training AI models on newspaper articles harms the market for and value of particular articles, it helps the market for and value of journalism.

Copyright is supposed to prevent unfair competition, but promote fair competition. The fourth fair use factor requires courts to determine whether the accused use of a copyrighted work represents fair or unfair competition. Their analysis should be informed by the first factor's transformativeness inquiry. *See, e.g., Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591 (1994) (holding that when the use of a work "is transformative, market substitution is at least less certain, and market harm may not be so readily inferred"). And their determination should ultimately depend on whether the accused use harms the market for or value of the copyrighted work in a way that affects the copyright owner's reasonable commercial expectations.

OpenAI's use of newspaper articles to train its generative AI models didn't harm the reasonable commercial expectations of the News Plaintiffs in any way. They commissioned and published newspaper articles with the expectation of selling consumers access to those articles. They retain and continue to exercise the exclusive right to sell consumers access to their articles. Their ability to sell consumers access to their articles wasn't affected byOpenAI's use of those articles to train its AI models.

C. Newspapers as Brands

The problem with the News Plaintiffs' fourth factor argument is that the defendant's use of their copyrighted work doesn't really affect the market for their primary product, which is their brand. While newspaper publishers have always been in the business of selling newspapers, which consist of newspaper articles, as a practical matter, newspaper publishers don't really sell newspaper articles, they sell a newspaper brand. Most consumers don't buy a copy of the *New York Times* in order to read a particular article. They buy a copy in order to read all the news that's fit to print, according to the *New York Times*.

What's more, news depends on novelty. As the Supreme Court has long recognized, "The value of the service, and of the news furnished, depends upon the promptness of transmission, as well as upon the accuracy and impartiality of the news, it being essential that the news be transmitted to members or subscribers as early or earlier than similar information can be furnished to competing newspapers by other news services." *International News Service v. Associated Press*, 248 U.S. 215, 230-31 (1918). Everyone wants the freshest news they can get, and it's a truism that "yesterday's news" is worthless. *Cf.* The Rolling Stones, *Yesterday's Papers* (1967) ("Who wants yesterday's papers? Nobody in the world.").

Of course, news articles may retain some residual economic value even after they become historical documents. But copyright doesn't entitle copyright owners to claim every last jot and tittle of economic value that might conceivably be derived from their works of authorship. The fair use doctrine quite explicitly provides that copyright doesn't give copyright owners the right to control many uses of their copyrighted works, including many economic uses. The fourth fair use factor doesn't prohibit uses that affect the market for or value of a work in any way, only those uses that "result in a substantially adverse

impact on the potential market for the original." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994) (internal quotation omitted). And when the use of a work "is transformative, market substitution is at least less certain, and market harm may not be so readily inferred." *Id.* at 591.

While factor four considers licensing markets, it only counts practical licensing markets. For example, in *American Geophysical Union v. Texaco*, 60 F.3d 913 (2d Cir. 1994), the court weighed harm to "traditional, reasonable, or likely to be developed" markets for the plaintiff's copyrighted works. It found that such a market existed based on the Copyright Clearance Center, an existing copyright clearinghouse that made licensing not only possible, but also practical. *See also Hachette v. Internet Archive*, 115 F.4th 163 (2d Cir. 2024) (finding harm to an existing e-book licensing market).

By contrast, no general licensing market for generative AI training on newspaper articles is practical or desirable. The primary market for news articles is exhausted when those articles are published and some are consumed as news. For better or worse, copyright gives newspapers the ability to control the use of their articles long after they cease to be news. But it doesn't and shouldn't give newspapers the right to control the use of their articles for transformative purposes that don't involve consumption and generate economic value. While some AI companies have licensed copyrighted works for the purpose of training their generative AI models, there is no general licensing market for using copyrighted works to train generative AI models, let alone a licensing market specific to copyrighted newspaper articles. And there is no reason to create one. On the contrary, using newspaper articles to train a generative AI model should be a fair use, so long as the model attributes its output to the newspaper articles on which it was trained.

II. The Economic Value of Attribution

The News Plaintiffs argue that the defendant has harmed the market for and value of their copyrighted works by diluting the market for newspaper articles. But their argument has no basis in copyright law. In an unpublished report, the United States Copyright Office suggested that courts could consider treating market dilution as a kind of market harm for the purpose of the infringement analysis, while acknowledging that no precedent for doing so exists. *See* United States Copyright Office, *Copyright and Artificial Intelligence Part 3: Generative AI Training* (Pre-Publication Version) (May 2025). And in *Kadrey v. Meta*, Judge Chhabria opined in dicta that the plaintiffs should have alleged dilution, without providing any authority to support a dilution to theory of harm. 788 F. Supp. 3d 1026 (N.D. Ca. 2025). Neither of these is even a persuasive authority for a dilution theory of harm, because neither provides any authority to substantiate such a theory. There is no authority to support a dilution theory of copyright harm, because dilution is not a copyright harm.

What's more, even if AI-generated works dilute the market for the human-generated works on which AI models are trained, AI companies can mitigate that harm by requiring AI models to credit the copyrighted works from which they drew information. Newspaper articles aren't only valuable because of copyright. They're also valuable because of attribution. While many authors value the copyright in their original works of authorship, most also value the attribution of those works, and many authors value attribution more than copyright. In fact, many authors only value attribution, and only use copyright to ensure attribution. That's why Creative Commons created its licenses, in order to enable authors to use copyright

to allow the widespread reproduction and distribution of their works, while requiring attribution of those works.

The paradigmatic Creative Commons license is the CC-BY or attribution license, which allows anyone to use the licensed work in any way, so long as they attribute the work to its author. See Creative Commons, CC-BY Attribution 4.0 International, at https://creativecommons.org/licenses/by/4.0/deed.en. Authors use the CC-BY license when they don't value the exclusive rights of copyright owners and are actually frustrated by the transaction costs copyright creates. Authors adopting Creative Commons licenses want anyone and everyone to use their works and circulate them as widely as possible. All they ask for in exchange is attribution, which copyright ironically doesn't itself require. In other words, Creative Commons licenses enable authors to use copyright to get recognition, which is what they really want but copyright doesn't necessarily provide.

Scholars often use Creative Commons licenses because they care about attribution, rather than copyright. As Oscar Wilde famously observed, the only thing worse than being talked about is not being talked about, and scholars are often ignored, especially when their works aren't conveniently available to the public. Creative Commons licenses enable scholars to make their works maximally available, by requiring nothing more than attribution, which most users of those works gladly provide.

Many scholars would welcome AI companies training AI models on their scholarship, so long as those models provided attribution. For example, some prominent legal scholars have argued that the legal academy should require the attribution of the ideas incorporated into AI-generated works. *See*, *e.g.*, Mark A. Lemley & Lisa Larrimore Ouellette, *Plagiarism, Copyright, and AI*, Chi. L. Rev. Online (Oct. 24, 2025), at https://lawreview.uchicago.edu/online-archive/plagiarism-copyright-and-ai. *But see* Brian L. Frye, *Aspirational Attribution: A Response to Lemley & Ouellette,* Plagiarism, Copyright, and AI, Chi. L. Rev. Online (Feb. 20, 2026), at https://lawreview.uchicago.edu/online-archive/aspirational-attribution-response-lemley-ouellette-plagiarism-copyright-and-ai (arguing that attribution should be voluntary, not mandatory).

The same is true - and has always been true - of newspapers and journalists. The primary value of news is in the reporting and being identified as the source of the reporting. With firm attribution shared across LLMs of which article or reporter has delivered a disclosure – adding to the pool of knowledge – incentives and concomitant rewards would increase for original, deep-dive reporting that is accelerated by "find the needle in a haystack" processing power of AI. As Justice Holmes famously observed:

> If a given person is to be prohibited from making the use of words that his neighbors are free to make, some other ground must be found. One such ground is vaguely expressed in the phrase unfair trade. This means that the words are repeated by a competitor in business in such a way as to convey a misrepresentation that materially injures the person who first used them, by appropriating credit of some kind which the first user has earned.

International News Service v. Associated Press, 248 U.S. 215, 247 (1918) (Holmes, J., concurring).

If training a generative AI model on copyrighted newspaper articles harms the market for or value of those articles by diluting the market for them, the harm can be mitigated by attribution. In fact, attributing the sources used by a generative AI model would fully mitigate any harm caused by dilution, especially in relation to newspapers and journalists. After all, the harm allegedly caused by dilution is that consumers won't attribute the news they consume with its original source, but instead will attribute it to a generative AI model. Automatic attribution by generative AI models would eliminate that harm and ensure that consumers continue to attribute the news to its original source, or at least to whatever source was used in training the AI model.

Generative AI models, with modest changes, are capable of tightly attributing information to sources. *See* Jack Hardinges & Irina Bejan, *Attribution in the Age of AI*, Creative Commons, at https://creativecommons.org/2026/07/24/attribution-in-the-age-of-ai/?utm_source=chatgpt.com. IIndeed, many already provide attribution. For example, I gave ChatGPT 5.6 Luna, one of OpenAI's generative AI models, the following prompt: "Please provide a brief explanation of the defendant's argument in New York Times v. OpenAI, focusing on the fourth fair use factor and providing citations to relevant sources." It provided the following response:

> OpenAI's argument under the fourth fair-use factor—the effect on the potential market—is that training GPT on Times articles does not significantly harm the market for those articles because the training process serves a different function: it converts text into statistical relationships used to generate predictions, rather than distributing the articles to readers. [1]
>
> More specifically, OpenAI argues that the relevant inquiry is whether the challenged use creates substitutive competition for the copyrighted works themselves. Relying heavily on Authors Guild v. Google, the defense maintains that ChatGPT does not provide the public with access to the Times articles used for training, so the training copies do not substitute for the articles in the legally relevant sense. [2] OpenAI also characterizes instances in which ChatGPT produces near-verbatim passages as anomalous memorization or a technical "bug," not the ordinary purpose or commercial function of the model. [3]
>
> The defense therefore separates training from allegedly infringing outputs: even if particular outputs reproduce protected expression or reduce traffic to Times content, OpenAI contends that those output-specific issues should not establish that the underlying training use causes cognizable market harm. [1] The Times, by contrast, argues that AI-generated answers can substitute for its journalism and that an existing market for licensing news content demonstrates lost licensing revenue; recent reporting describes evidence that Microsoft executives acknowledged reduced click-throughs and substitution by chatbot answers. [4]
>
> Relevant authorities include Authors Guild v. Google (the principal precedent invoked by the defense), the statutory market-effect factor in 17 U.S.C. § 107, and the parties' competing summary-judgment arguments in the consolidated litigation. [2] [3]

[1] https://www.datastudios.org/post/us-government-backs-openai-new-york-times-copyright-fair-use-ai-training
[2] https://stanfordtechreview.com/articles/doj-ai-training-fair-use-openai-brief
[3] https://ailawsuittracker.com/cases/new-york-times-v-openai/
[4] https://digiday.com/media/paywall-violation-eviscerates-fair-use-defense-what-openai-nyt-twist-means-for-publishers/

As is apparent from this example, ChatGPT is capable of providing attribution to the sources on which it was trained, and already does so, at least to a limited degree. Newspapers and journalists value attribution, at least as much as, if not more than, copyright. AI companies can adequately compensate them for the reasonable residual market value of their news articles by giving them attribution. Doing so would benefit newspapers and journalists in the coin they value the most, without costing AI companies anything or harming the functionality of their generative AI models in any way. In fact, providing additional attribution would probably even improve the functionality and value of those generative AI models. Ironically, one of the reasons AI models don't currently provide more attribution is probably the risk of copyright infringement liability. If generative AI models routinely identified their sources, it would make it easier for the copyright owners of those sources to allege infringement. But if courts conclude that attribution is sufficient to mitigate any harm caused by dilution, then AI companies will have no reason to avoid attribution. On the contrary, they will have an incentive to attribute as extensively as possible, in order to reduce their risk of copyright infringement liability.

III. The Anti-Competitive Effects of Requiring Licensing

If courts determine that training a generative AI model on copyrighted newspaper articles is not a fair use, it will harm consumers by reducing competition in the market for newspaper articles in two different ways. First, it will harm consumers directly, by increasing the cost of newspaper articles. According to the News Plaintiffs, generative AI models make it faster, easier, and cheaper to produce newspaper articles, thereby increasing the supply. The law of supply and demand dictates that if the supply increases, the price will decrease. So reducing or eliminating access to generative AI models trained on newspaper articles is sure to make newspaper articles more costly. Second, it will harm consumers indirectly, by decreasing competition among AI companies. If AI companies are forced to license newspaper articles in order to train their generative AI models, only the largest and wealthiest AI companies will be able to compete. The AI market is already highly stratified, and requiring licensing would irrevocably cement that stratification.

While OpenAI is the defendant in this copyright infringement action, it may be agnostic about the outcome. If OpenAI wins, it will obviously be pleased to avoid paying licensing fees to the copyright owners of its training data. But none of its competitors will have to pay licensing fees either. As a consequence, it will be much easier for them to improve their own generative AI models and compete with OpenAI. While OpenAI is a leader of the pack for the time being, many hungry competitors are nipping at its heels.

As a consequence, OpenAI may welcome a holding that training a generative AI model is an infringing use of a copyright work that requires a license. After all, with a market value of about $852 billion, OpenAI can afford to license anything it wants to use as training data. But many of its potential competitors can't, and will fold if confronted with an obligation to obtain licenses for training data. Even worse, OpenAI could insist on exclusive licenses and shut out its competitors entirely.

This outcome would have highly anti-competitive effects. The market for generative AI models is already dominated by a few, highly capitalized companies. Requiring copyright licenses to train a generative AI model on copyrighted works would further consolidate their market dominance, because the companies currently dominating the market can afford to pay for copyright licenses, but their smaller competitors cannot.

Conclusion

The fair use inquiry is no mere doctrinal exercise. It requires - obligates - the court to consider the public policy implications of finding that a particular use of a copyrighted work is infringing. Obviously, the swarm of copyright infringement actions against AI companies for training their AI models on copyrighted works are immensely significant. In considering OpenAI's fair use defense, this court should consider the consequences of its conclusion. If training generative AI models on copyrighted works isn't a fair use, a few large copyright owners will receive an immense windfall, at crippling expense to the now-vibrant United States AI industry and to the consumers who benefit from AI technology.

As the Supreme Court has observed, courts should be reluctant to find that promising new technologies infringe copyright, especially when copyright owners never contemplated such a use of their works:

> Sound policy, as well as history, supports our consistent deference to Congress when major technological innovations alter the market for copyrighted materials. Congress has the constitutional authority and the institutional ability to accommodate fully the varied permutations of competing interests that are inevitably implicated by such new technology.

Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 431 (1984). In *Sony*, the Supreme Court found that fair use protected video tape recorders, an important new technology that promised considerable benefits to consumers. In the event, it also ended up creating an entirely new and lucrative market for the very copyright owners who opposed it. Generative AI is obviously an even more important technology than video tape recorders. It has already delivered immense economic and social value, and promises to transform just about every aspect of the modern world. As the Supreme Court observed in Sony, the courts should leave it to Congress to decide whether and how to regulate this revolutionary new technology.

The News Plaintiffs and the defendants are caught in a prisoner's dilemma. Both sides would be better off cooperating, but neither can agree to cooperate, for fear their opponent will defect. The News Plaintiffs could and should simply allow AI companies to train generative AI models on their newspaper articles in exchange for attribution, which – with some slight restructuring by AI companies of the generative presentation – drive traffic and increase readership and profits. And the defendants should offer to fully

and robustly attribute the output of their generative AI models to the newspaper articles on which they trained those models, with emphasis on those who add to the pool of knowledge with disclosures, scoops, new insights and breaking news. But neither side can propose this compromise, for fear of ruining their litigation position. The News Plaintiffs would have to admit that generative AI training doesn't really harm the market for or value of their works. And the defendants would have to admit that they are training their generative AI models on copyrighted newspaper articles.

This action is haunted by *Authors Guild v. Google*, 804 F.3d 202 (2nd Cir. 2015), which both sides claim supports their position. Just like this action, *Authors Guild v. Google* reflected a prisoner's dilemma. The Second Circuit tried to provide a solomonic solution, but ended up cutting the baby in half, because it forgot that the public was also a party in interest. A decade later, the public still doesn't have full access to the vast collection of books scanned by Google, even though almost all of those books have no commercial value and no one objects to Google making them available to the public.

This court can solve the market failure the prisoner's dilemma is causing in this action by helping the News Plaintiffs and the defendants cooperate and reach a mutually beneficial outcome that will also benefit the public. The News Plaintiffs have a right to profit from their copyrighted newspaper articles and the defendants have a right to train their generative AI models. But the public also has a right to use generative AI models and access the information they provide. After all, the whole point of copyright is to benefit the public.

Dated: October 27, 2026
By: /s/ Brian L. Frye
Brian L. Frye (BF1028)