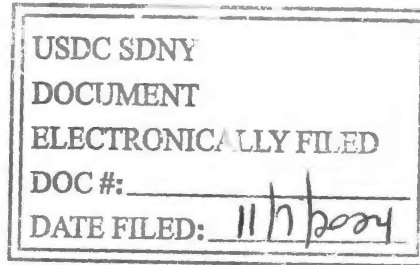


UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK



_____ X

RAW STORY MEDIA, Inc., ALTERNET MEDIA, INC.,

Plaintiffs,

vs.

24 Civ. 01514

OPENAI, INC., OPENAI GP, LLC,
OPENAI, LLC, OPENAI OPCO LLC,
OPENAI GLOBAL LLC, and OPENAI
HOLDINGS, LLC,

Defendants.

_____ X

DECISION AND ORDER

McMahon, J.:

Plaintiffs Raw Story Media, Inc. and AlterNet Media, Inc. (collectively “Plaintiffs”) brought this action, pursuant to the Digital Millennium Copyright Act (the “DMCA”), 17 U.S.C. § 1201 *et seq.*, against OpenAI, Inc., OpenAI GP, LLC, OpenAI, LLC, OpenAI Opco LLC, OpenAI Global LLC, and OpenAI Holdings, LLC (collectively, “Defendants” or “OpenAI”). OpenAI seeks to dismiss Plaintiffs’ complaint in its entirety, pursuant to Fed. R. Civ. P. 12(b)(1) and 12(b)(6). Dkt. No. 68 (“MTD”).

For the reasons below, OpenAI’s motion to dismiss is GRANTED. Plaintiffs’ motion for leave to replead is DENIED WITHOUT PREJUDICE to renewal on a proper record—which means filing a notice of motion to which a proposed amended pleading is attached, together with an explanation of why the proposed amendment would not be futile.

BACKGROUND

Plaintiffs Raw Story Media, Inc. and AlterNet Media, Inc. are news organizations that have published, collectively, “more than 400,000 breaking news features, investigative news articles and opinion columns” online. Dkt. No. 1, ¶ 14 (“Compl.”). Defendants OpenAI are seven interrelated organizations, doing business within the State of New York, responsible for an AI service known as ChatGPT. Compl. ¶¶ 15-22, 24. ChatGPT is an AI-powered large language model (“LLM”) that allows paying users to enter text prompts to which ChatGPT will generate responses. According to Plaintiffs, ChatGPT “gives the impression that it is an all-knowing ‘intelligent’ source of the information being provided.” Compl. ¶ 40. However, “ChatGPT does not have any independent knowledge of the information provided in its responses.” Compl. ¶ 36. Rather, ChatGPT is trained on large amounts of text, known as “training sets.” Compl. ¶ 29. “These training sets range from collections of links posted on the website Reddit to a scrape of most of the internet.” Compl. ¶ 29.

Plaintiffs allege that “thousands” of their copyrighted works of journalism were caught in this “scrape,” stripped of their author, title, and copyright information, and input into at least three of OpenAI’s training sets (WebText, WebText2, and Common Crawl). Compl. ¶¶ 29, 30, 37. Plaintiffs allege that these three training sets were then used to train ChatGPT. Compl. ¶ 29. Since ChatGPT was not provided with the author, title, and copyright information, Plaintiffs allege that ChatGPT would not have learned to communicate that information when fashioning responses to user inquiries that are based on their copyrighted works, Compl. ¶ 38, and that indeed ChatGPT “generally does not provide the author, title, and copyright information applicable to the works on which its responses are based.” Compl. ¶ 39. Plaintiffs allege that Defendants’ removal of copyright management information (“CMI”) from Plaintiffs’ works—prior to training ChatGPT—

is a violation of Section 1202(b)(i) of the Digital Millennium Copyright Act (the “DMCA”), for which Plaintiffs are entitled to actual or statutory damages.

Plaintiffs further seek injunctive relief against Defendants. Plaintiffs allege that earlier versions of ChatGPT generated significant amounts of plagiarized content. Compl. ¶ 5. If Plaintiffs’ works remain in ChatGPT’s repository without any CMI, Plaintiffs allege that there is a substantial likelihood that the current version of ChatGPT will reproduce, verbatim or nearly verbatim, Plaintiffs’ copyrighted works without providing the author, title, or copyright information contained in those works. *See* Prayer for Relief.

OpenAI has moved to dismiss the complaint. Dkt. No. 68 (“MTD”). Plaintiff opposes the motion. Dkt. No. 70 (“Opp.”).

LEGAL STANDARDS

I. Section 1202(b)(i) of the DMCA

Section 1202(b)(i) of the DMCA provides that:

No person shall, without the authority of the copyright owner or the law,... intentionally remove or alter any [CMI]...knowing, or with respect to civil remedies under section 1203, having reasonable grounds to know that it will induce enable, facilitate or conceal an infringement of any right under this title.

17 U.S.C. §1202(b). The Second Circuit has found this statute to incorporate a so-called double-scienter requirement: that the defendant both know “that CMI has been removed or altered without authority of the copyright owner or the law” and know, “or hav[e] reasonable grounds to know that such distribution will induce, enable, facilitate, or conceal an infringement.” *Mango v. Buzzfeed*, 970 F.3d 167, 171 (2d Cir. 2020) (cleaned up) (internal citations omitted).

II. Article III Standing

“Article III standing requires a concrete injury even in the context of a statutory violation.” *Spokeo, Inc. v. Robins*, 578 U.S. 330, 341 (2016). “[T]o establish standing, a plaintiff must show (i) that he suffered an injury in fact that is concrete, particularized, and actual or imminent; (ii) that the injury was likely caused by the defendant, and (iii) that the injury would likely be redressed by judicial relief.” *TransUnion LLC v. Ramirez*, 594 U.S. 413, 423 (2021). “The party invoking federal jurisdiction bears the burden of establishing these elements.” *Lujan v. Defenders of Wildlife*, 504 U.S. 555, 561 (1992).

“Where, as here, a case is at the pleading stage, the plaintiff must clearly...allege facts demonstrating each element.” *Spokeo, Inc. v. Robins*, 578 U.S. 330, 338 (2016) (internal citation and quotations omitted). “[S]tanding is not dispensed in gross; rather, plaintiffs must demonstrate standing...for each form of relief that they seek (for example, injunctive relief and damages).” *TransUnion LLC*, 594 U.S. at 431.

“What makes a harm concrete for the purposes of Article III?” *TransUnion*, 594 U.S. at 424:

As a general matter, the Court has explained that history and tradition offer a meaningful guide to the types of cases that Article III empowers federal courts to consider. And with respect to the concrete-harm requirement in particular, this Court's opinion in *Spokeo v. Robins* indicated that courts should assess whether the alleged injury to the plaintiff has a close relationship to a harm traditionally recognized as providing a basis for a lawsuit in American courts. That inquiry asks whether plaintiffs have identified a close historical or common-law analogue for their asserted injury. *Spokeo* does not require an exact duplicate in American history and tradition. But *Spokeo* is not an open-ended invitation for federal courts to loosen Article III based on contemporary, evolving beliefs about what kinds of suits should be heard in federal courts.

Id. at 424-25. (internal quotations and citations omitted).

DISCUSSION

Defendants seek to dismiss Plaintiffs' complaint in its entirety. Defendants argue that Plaintiffs lack Article III standing to assert their claims, and that the Court therefore lacks subject-matter jurisdiction over this action pursuant to Fed. R. Civ. P. 12(b)(1). Defendants argue, in the alternative, that Plaintiffs have failed to state a claim upon which relief can be granted, pursuant to Fed. R. Civ. P. 12(b)(6).

“Whether a claimant has standing is ‘the threshold question in every federal case, determining the power of the court to entertain the suit.’” *U.S. v. Cambio Exacto, S.A.*, 166 F.3d 522, 526 (2d Cir. 1999) (quoting *In re Gucci*, 126 F.3d 380, 387-88 (2d Cir.1997)). Accordingly, I begin my inquiry here.

Plaintiffs argue that they have standing to pursue two forms of relief. *First*, Plaintiffs argue that they have standing to pursue damages because “the unlawful removal of CMI from a copyrighted work is a concrete injury.” *Opp.* at 7. *Second*, Plaintiffs argue that they have standing to pursue injunctive relief, because they have alleged that there is a substantial risk that Defendants' program will “provide responses to users that incorporate[] material from Plaintiffs' copyright-protected works or regurgitate[] copyright-protected works verbatim or nearly verbatim.” *Compl.* ¶ 52; *see also Opp.*, at 9-10. Defendants respond that neither theory of harm identifies a concrete injury-in-fact sufficient to establish standing.

I agree with Defendants. Plaintiffs' claims for both damages and injunctive relief are DISMISSED because Plaintiffs' lack Article III standing. Accordingly, I need not reach the alternative motion pursuant to Fed. R. Civ. P. 12(b)(6).

I. Plaintiffs lack standing to pursue their claim for damages

Let us consider Plaintiffs' contention in respect of their claim for damages: that the unauthorized removal of CMI from their copyrighted work gives rise to a concrete injury-in-fact, even though they do not allege that a copy of their work from which the CMI has been removed has been disseminated by ChatGPT to anyone in response to any specific query. *See* Opp., at 7. Plaintiffs contend that their injury bears a "close relationship" to the tort of copyright infringement, because "the protection against removing or altering CMI, 17 U.S.C. § 1202(b)(i), is analogous to the rights to reproduce the works and prepare derivative ones, 17 U.S.C. §§ 106(1), (2)" in that "both grant the copyright owner the sole prerogative to decide how future iterations of the work may differ from the version the owner published." Opp., at 6. Plaintiffs argue that this, in turn, "accords with the common law, which recognizes interference with property, without more, as a concrete injury." *Id.* at 7.

I am not convinced that injury for interference with property provides the necessary "close historical or common-law analogue" to Plaintiffs' alleged injury. For one thing, Plaintiffs are wrong that Section 1202 "grant[s] the copyright owner the sole prerogative to decide how future iterations of the work may differ from the version the owner published." Other provisions of the Copyright Act afford such protections, *see* 17 U.S.C. § 106, but not Section 1202. Section 1202 protects copyright owners from specified interferences with the *integrity of a work's CMI*. In other words, Defendants may, absent permission, reproduce or even create derivatives of Plaintiffs' works—*without incurring liability under Section 1202*—as long as Defendants keep Plaintiffs' CMI intact. Indeed, the legislative history of the DMCA indicates that the Act's purpose was not to guard against property-based injury. Rather, it was to "ensure the integrity of the electronic marketplace by preventing fraud and misinformation," and to bring the United States into

compliance with its obligations to do so under the World Intellectual Property Organization (WIPO) Copyright Treaty, art. 12(1) (“Obligations concerning Rights Management Information”) and WIPO Performances and Phonograms Treaty, art. 19 (same). H.R. Rep. No. 105-551, at 10-11.

Moreover, I am not convinced that the mere removal of identifying information from a copyrighted work—absent dissemination—has *any* historical or common-law analogue.

TransUnion is clear: “the plaintiff’s injury [must] in fact be concrete—that is, real and not abstract. 594 U.S. at 424 (internal quotations and citations omitted). Plaintiffs allege that their copyrighted works (absent CMI) were used to train an AI-software program and remain in ChatGPT’s repository of text. But Plaintiffs have not alleged any *actual* adverse effects stemming from this alleged DMCA violation. The argument advanced by Plaintiffs is akin to that of the *dissent* in *TransUnion*: “If a [defendant] breaches a [DMCA] duty owed to a specific [copyright owner], then that [copyright owner]...[has] a sufficient injury to sue in federal court.” *Id.* at 450 (Thomas, J., dissenting). To this, the majority of the Court said: ‘no.’ “No concrete harm, no standing.” *Id.* at 442. Accordingly, Plaintiffs lack Article III standing to seek retrospective relief in the form of damages for the injury they allege.

A. Plaintiffs lack standing to pursue their claim for injunctive relief

And so we turn to Plaintiffs’ claim for injunctive relief. Plaintiffs seek an injunction “requiring Defendants to remove all copies of Plaintiffs’ copyrighted works from which author, title, copyright, and terms of use information w[ere] removed from their training sets and any other repositories.” Prayer for Relief.

Plaintiffs argue that they are entitled to such an injunction because, whether ChatGPT has or has not already reproduced their copyrighted work without attaching the required CMI, there is a substantial risk that ChatGPT will do so in the future. “[A] person exposed to a risk of future harm may pursue forward-looking, injunctive relief to prevent the harm from occurring, at least so long as the risk of harm is sufficiently imminent and substantial.” *TransUnion*, 594 U.S. at 435. “An allegation of future injury may suffice if the threatened injury is ‘certainly impending,’ or there is a ‘substantial risk that the harm will occur.’” *Susan B. Anthony List v. Driehaus*, 573 U.S. 149, 158 (2014) (quoting *Clapper v. Amnesty Int’l USA*, 568 U.S. 398, 414 n.5 (2013)). A substantial risk means there is a “realistic danger of sustaining a direct injury.” *Pennell v. City of San Jose*, 485 U.S. 1, 8 (1988) (quoting *Babbitt v. United Farm Workers Nat’l Union*, 442 U.S. 289, 298 (1979)).

Defendants concede that there are clear historical and common-law analogues for this type of injury. Dkt. No. 71 (“Reply”), at 5 (citing *Masson v. New Yorker Mag., Inc.*, 501 U.S. 496, 511-12 (1991) (defamation); *Petrone v. Turner Publ’n Co. LLC*, 2023 WL 7302447, at *4 (S.D.N.Y. Nov. 6, 2023) (false light); *Gilliam v. Am. Broad. Cos., Inc.*, 538 F.2d 14, 24 (2d Cir. 1976) (unfair competition); *Lamothe v. Atl. Recording Corp.*, 847 F.2d 1403, 1406-07 (9th Cir. 1988) (“reverse passing off”); *Macia v. Microsoft Corp.*, 152 F. Supp. 2d 535, 541 (D. Vt. 2001) (slander of title)). However, Defendants argue that Plaintiffs lack standing to seek injunctive relief because they fail to allege facts tending to show that the risk of ChatGPT reproducing Plaintiffs’ work, in whole or in part, absent the requisite CMI is “substantial.” *See Reply*, at 3.

I agree with Defendants. Plaintiffs allege that ChatGPT has been trained on “a scrape of most of the internet,” Compl. ¶ 29, which includes massive amounts of information from innumerable sources on almost any given subject. Plaintiffs have nowhere alleged that the

information in their articles is copyrighted, nor could they do so. When a user inputs a question into ChatGPT, ChatGPT synthesizes the relevant information in its repository into an answer. Given the quantity of information contained in the repository, the likelihood that ChatGPT would output plagiarized content from one of Plaintiffs' articles seems remote. And while Plaintiffs provide third-party statistics indicating that an *earlier* version of ChatGPT generated responses containing significant amounts of plagiarized content, Compl. ¶ 5, Plaintiffs have not plausibly alleged that there is a "substantial risk" that the *current* version of ChatGPT will generate a response plagiarizing one of *Plaintiffs'* articles.

Accordingly, Plaintiffs lack Article III standing to seek injunctive relief for their alleged injury.

* * *

Let us be clear about what is really at stake here. The alleged injury for which Plaintiffs truly seek redress is not the exclusion of CMI from Defendants' training sets, but rather Defendants' *use* of Plaintiffs' articles to develop ChatGPT without compensation to Plaintiffs. *See* Compl. ¶ 57 ("The OpenAI Defendants have acknowledged that use of copyright-protected works to train ChatGPT requires a license to that content, and in some instances, have entered licensing agreements with large copyright owners...They are also in licensing talks with other copyright owners in the news industry, but have offered no compensation to Plaintiffs."). Whether or not that type of injury satisfies the injury-in-fact requirement, it is not the type of harm that has been "elevated" by Section 1202(b)(i) of the DMCA. *See Spokeo*, 578 U.S. at 341 (Congress may "elevate to the status of legally cognizable injuries, *de facto* injuries that were previously inadequate in law."). Whether there is another statute or legal theory that does elevate this type of harm remains to be seen. But that question is not before the Court today.

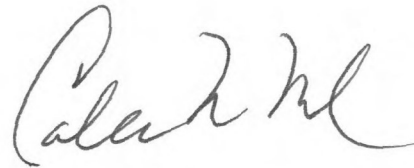
In the event of dismissal Plaintiffs seek leave to file an amended complaint. I cannot ascertain whether amendment would be futile without seeing a proposed amended pleading. I am skeptical about Plaintiffs' ability to allege a cognizable injury but, at least as to injunctive relief, I am prepared to consider an amended pleading.

Conclusion

For the foregoing reasons, Defendants' motion to dismiss is GRANTED in its entirety. Plaintiffs' motion for leave to replead is DENIED WITHOUT PREJUDICE to renewal on a proper record—which means filing a notice of motion to which a proposed amended pleading is attached together with an explanation of why the proposed amendment would not be futile.

This constitutes a written opinion. The Clerk is directed to remove the motion at Dkt. No. 68 from the Court's list of open motions.

Dated: November 7, 2024

A handwritten signature in cursive script, appearing to read "Caleb M.", written above a horizontal line.

U.S.D.J.

BY ECF TO ALL COUNSEL