UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

-- X
:
RAW STORY MEDIA, INC., ALTERNET :   NO. 1:24-cv-01514 CM
MEDIA, INC., :
:
:
Plaintiffs, :
:
:
v. :
:
:
OPENAI, INC., OPENAI GP, LLC, :
OPENAI, LLC, OPENAI OPCO LLC, :
OPENAI GLOBAL LLC, OAI :
CORPORATION, LLC, OPENAI :
HOLDINGS, LLC, :
:
:
Defendants. :
:
:
-- X

OPENAI'S MEMORANDUM OF LAW IN OPPOSITION TO PLAINTIFFS' MOTION
FOR LEAVE TO AMEND COMPLAINT OR, IN THE ALTERNATIVE, TO CONTINUE
TAKING JURISDICTIONAL DISCOVERY

TABLE OF CONTENTS

Page

TABLE OF CONTENTS
(continued)

TABLE OF AUTHORITIES

Page(s)

Cases

TABLE OF AUTHORITIES
(continued)

TABLE OF AUTHORITIES
(continued)

Page(s)

I.   INTRODUCTION

"No concrete harm, no standing." Dkt. No. 117 ("Ord.") at 7 (quoting *TransUnion LLC v. Ramirez*, 594 U.S. 413, 442 (2021)). When this Court dismissed Plaintiffs' complaint, it concluded that Plaintiffs' asserted harm—*i.e.*, "the unauthorized removal of CMI from their copyrighted work[,] even though they do not allege that a copy of their work from which the CMI has been removed has been disseminated by ChatGPT"—was insufficiently concrete. Ord. at 6. The Court was "skeptical about Plaintiffs' ability to allege a cognizable injury," but indicated willingness to consider an amended pleading "at least as to injunctive relief," to the extent Plaintiffs could allege "a substantial risk that the *current* version of ChatGPT will generate a response plagiarizing one of *Plaintiffs'* articles." Ord. at 8-10 (emphasis in original). Plaintiffs have now proposed an amended pleading, but it does not meet that standard. Instead, Plaintiffs reallege the *exact same* harm that the Court already rejected—that OpenAI "injured Plaintiffs by removing CMI from tens of thousands of Plaintiffs' news articles." Dkt. No. 119 ("Mot.") at 1.

Plaintiffs' motion for leave to amend spills much ink relitigating arguments already lost, including by trying to draw a new common law analogy that relies on the same faulty allegation of harm. But Plaintiffs still fail to allege that ChatGPT has ever disseminated its works without CMI, nor do they allege a substantial risk that ChatGPT will imminently do so. Accordingly, Plaintiffs' proposed amendments are futile, and their motion for leave to amend should be denied.

II.   BACKGROUND

Plaintiffs filed this action on February 28, 2024, alleging that OpenAI "created copies of Plaintiffs' works of journalism . . . and included them in training sets used to train ChatGPT" with author, title, and copyright notice information removed in violation of 17 U.S.C. § 1202(b)(1). Dkt. No. 1 ("Compl.") ¶¶ 49–51. OpenAI moved to dismiss the complaint on April 29, 2024.

1

On November 7, 2024, this Court granted OpenAI's motion to dismiss the complaint in its entirety on Article III standing grounds. Ord. at 5. The Court concluded that, absent allegations of dissemination, "injury for interference with property" did not "provide[] the necessary close historical or common-law analogue to Plaintiffs' alleged injury." Ord. at 6. This was because Section 1202(b)'s "purpose was not to guard against property-based injury," but instead "to ensure the integrity of the electronic marketplace by prevent fraud and misinformation, and to bring the United States into compliance with its obligations to do so" under international treaties. *Id.* at 6–7. The Court was "not convinced that the mere removal of identifying information from a copyrighted work—absent dissemination—has *any* historical or common law analogue." *Id.* at 7 (emphasis in original). After all, Plaintiffs had not "alleged any *actual* adverse effects stemming from this alleged DMCA violation." *Id.* (emphasis in original). Plaintiffs' argument was "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.* (formatting in original).

The Court also found that Plaintiffs lacked standing to pursue injunctive relief because "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." Ord. at 9 (emphasis in original). Because Plaintiffs alleged that "ChatGPT has been trained on a scrape of most of the internet, which includes massive amounts of information from innumerable sources on almost any given subject, . . . the likelihood that ChatGPT would output plagiarized content from one of Plaintiffs' articles seems remote." Ord. at 8–9).

Ultimately, while the Court was "skeptical about Plaintiffs' ability to allege a cognizable injury" absent allegations of dissemination, it was "prepared to consider an amended pleading,"

2

"at least as to injunctive relief." Ord. at 10. Plaintiffs filed a motion for leave to file an amended complaint on November 21, 2024. *See* Mot. (Dkt. No. 119) with Proposed First Amended Complaint ("PFAC"), Dkt. No. 119-1 ("PFAC"). OpenAI opposes.

III.   LEGAL STANDARD

While "[t]he court should freely give leave when justice so requires," Fed. R. Civ. P. 15(a)(2), the Second Circuit has emphasized that leave to amend should not be granted automatically, particularly where the proposed amendment would be futile. *McCarthy v. Dun & Bradstreet Corp.*, 482 F.3d 184, 200 (2d Cir. 2007). A proposed amended complaint is futile if it does not "state[] sufficient facts to survive a motion to dismiss." *Keller v. Schoharie Cty. Dept. of Social Servs.*, 848 F. App'x 38, 39 (2d Cir. 2021).

IV.   ARGUMENT

A.   Plaintiffs' proposed amendments are futile.

Plaintiffs' allegations are still insufficient to establish standing under Rule 12(b)(1). Plaintiffs lack standing to pursue injunctive relief because they still "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." *See* Ord. at 9 (emphasis in original). Plaintiffs also lack standing to pursue damages: they admit they cannot allege dissemination (PFAC ¶ 67), and their attempts to relitigate their failed analogy or draw a new one to unjust enrichment are unavailing. In any event, Plaintiffs' Proposed First Amended Complaint cannot survive a Rule 12(b)(6) motion to dismiss on the merits, either. Plaintiffs' proposed amendments are futile, and their motion for leave to amend should be denied.

1.   Plaintiffs' proposed amendments do not establish standing to pursue injunctive relief.

In its Order dismissing the complaint, this Court articulated a standard for establishing standing to pursue injunctive relief: Plaintiffs must "plausibly allege[] that there is a 'substantial risk' that the *current version* of ChatGPT will generate a response plagiarizing one of *Plaintiffs'* articles." Ord. at 9 (emphasis in original). Plaintiffs' proposed amendments do not meet this standard.

Plaintiffs now admit that they tried—and failed—to force ChatGPT to disseminate one of their works. *See* PFAC ¶ 67. Setting aside that this "would ordinarily be cause for celebration, not a lawsuit," *TransUnion*, 594 U.S. at 437, Plaintiffs' admission is fatal to their pursuit of injunctive relief. Far from a substantial risk that ChatGPT will output one of their articles, Plaintiffs' proposed amendments instead show the opposite: despite counsel's best efforts, Plaintiffs could not artificially manufacture a regurgitation. As Plaintiffs' own allegations show, the risk of a real-world dissemination occurring is not "sufficiently imminent and substantial" to show standing. *See* Ord. at 8 (quoting *TransUnion*, 594 U.S. at 435).

In an attempt to distract from their failure to manufacture a dissemination, Plaintiffs propose new allegations that ChatGPT has "regurgitated" or "abridged" *other* publishers' works. PFAC ¶¶ 65, 67–68. But Plaintiffs do not attempt to explain how these allegations demonstrate that a dissemination of one of *their own* articles is "certainly impending," or that Plaintiffs face a "realistic danger of" dissemination. *See* Ord. at 8 (quoting *Susan B. Anthony List v. Driehaus*, 573 U.S. 149, 158 (2014); *Pennell v. City of San Jose*, 485 U.S. 1, 8 (1988)).

Nor could they, as doing so would be "too speculative to support Article III standing." *TransUnion*, 594 U.S. at 438. In *TransUnion*, the subgroup of plaintiffs whose credit information was disseminated to potential creditors had standing, while those whose information was not

4

disseminated did not. *Id.* at 439. If Plaintiffs' argument were correct, *all* the *TransUnion* plaintiffs would have had standing simply because TransUnion disseminated at least some *other* person's credit report. But the Supreme Court concluded otherwise: even though "TransUnion could have divulged their misleading credit information to a third party at any moment," the plaintiffs whose reports were not disseminated "did not demonstrate a sufficient likelihood that their individual credit information would be requested by third-party businesses and provided by TransUnion during the relevant time period." *Id.* at 438. The same is true here. Allegations that ChatGPT has regurgitated or abridged *other* publishers' articles does not demonstrate a sufficient likelihood that ChatGPT will regurgitate or abridge one of *Plaintiffs'*.

Plaintiffs also fail to address a critical flaw that the Court identified in their original complaint: that Plaintiffs do not demonstrate a substantial risk of dissemination because Plaintiffs allege that OpenAI's models have been trained on "a scrape of most of the internet." PFAC ¶ 53; Compl. ¶ 29. Accordingly, and even viewing Plaintiffs' allegations in the light most favorable to them, "the likelihood that ChatGPT would output plagiarized content from one of Plaintiffs' articles seems remote." Ord. at 9. Rather than addressing this deficiency, Plaintiffs attempt to add detailed allegations about OpenAI's technical processes. *See generally* PFAC ¶¶ 31–62. But these details do not explain why *Plaintiffs'* articles, as opposed to any other of the "massive amounts of information from innumerable sources on almost any given subject," face a substantial risk of dissemination. *See* Ord. at 8. Because Plaintiffs have failed to plausibly allege a "substantial likelihood that the *current* version of ChatGPT will generate a response plagiarizing one of *Plaintiffs'* articles," Plaintiffs lack standing to pursue injunctive relief. *See* Ord. at 9 (emphasis in original).

2.   Plaintiffs' proposed amendments do not establish standing to pursue damages.

Plaintiffs likewise fail to plead standing to seek damages. Their proposed amendments are futile for at least two independent reasons. First, Plaintiffs' new factual allegations do not—and cannot—overcome the *legal* deficiencies that doomed Plaintiffs' first complaint. Nor do they warrant reconsideration of this Court's prior legal conclusion that absent dissemination, the removal of copyright management information is not a sufficiently concrete injury. Second, even if the Court elects to consider Plaintiffs' new legal theories, they still fail on the merits.

a.   Plaintiffs' new factual allegations cannot fix the legal shortcomings with Plaintiffs' alleged injury.

In its prior order, this Court found that "Plaintiffs lack Article III standing to seek retrospective relief in the form of damages for the injury they allege." Ord. at 7. The Court was "not convinced" that there was "*any* historical or common-law analogue" to Plaintiffs' asserted injury: "the mere removal of identifying information from a copyrighted work[,] absent dissemination." *Id.* (emphasis in original). Plaintiffs' proposed amendment does not cure that defect. It does not allege dissemination, nor any new injury. Instead, Plaintiffs double down and reassert the *same* alleged injury that the Court already rejected. The first sentence of their motion proclaims that "Defendants injured Plaintiffs by removing CMI from tens of thousands of Plaintiffs' news articles." Mot. at 1. Plaintiffs' argument thus confirms that the Proposed First Amended Complaint is based on the same harm—*i.e.*, "the mere removal of identifying information from a copyrighted work"—that the Court has deemed insufficient as a matter of law. *See* Ord. at 7.

Because the Court found this asserted harm to be insufficient as a matter of law, Plaintiffs' proposed amendments are futile. *See In re OSG Securities Litig.*, 62 F. Supp. 3d 353, 357–58 (S.D.N.Y. 2014) (denying motion for leave to amend where plaintiffs' new allegations were

"deficient as a matter of law"); *Liantonio v. Lavintman*, No. 11-cv-1292(SJF)(AKT), 2012 WL 4483040, at *5 (E.D.N.Y. Sept. 27, 2012) ("As each of the proposed amended claims fail as a matter of law, the motion for leave to amend is denied.").   The Court based its dismissal of Plaintiffs' claims for damages on its interpretation of Section 1202(b)'s text and legislative history, not on the sufficiency of Plaintiffs' factual allegations.   *See* Ord. at 6. Regarding Plaintiffs' copyright infringement analogy, the Court explained that "injury for interference with property" cannot provide the necessary close historical or common-law analogue because Section 1202 does not "grant the copyright owner the sole prerogative to decide how future iterations of the work may differ from the version the owner published." Ord. at 6.   Instead, the statute's purpose is 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)." Ord. at 6–7.   Because the problems with Plaintiffs' harm assertions are "substantive rather than the result of an inadequately or inartfully pleaded complaint, an opportunity to replead would be futile and should be denied." *In re Sanofi Sec. Litig.*, 87 F. Supp. 3d 510, 548–49 (S.D.N.Y. 2015).

Plaintiffs now attempt to raise a new analogy between their purported injury and unjust enrichment.   *See* Mot. at 6.   But this analogy, too, relies on the same underlying asserted harm: that Plaintiffs' "copyrighted works (absent CMI) were used to train an AI-software program and remain in ChatGPT's repository of text." *See* Ord. at 7.   Plaintiffs still "have not alleged any *actual* adverse effects stemming from [OpenAI's] alleged DMCA violation." Ord. at 7 (emphasis in original).   As with Plaintiffs' stubborn reassertion of a copyright infringement analogy, amendment is futile because the Court has already decided as a matter of law that standing requires dissemination of Plaintiffs' works without CMI.   *See* Ord. at 7.   Without allegations of

dissemination, or a new factual theory of how they were harmed, Plaintiffs lack Article III standing to pursue damages.

> **b.   Instead of alleging the necessary facts, Plaintiffs simply quarrel with the Court's legal conclusions.**

In any event, because Plaintiffs' proposed amended complaint is premised on the same asserted injury as the original, Plaintiffs' motion for leave is a motion for reconsideration in disguise. *See Dube v. Signet Jewelers Ltd.*, No. 16-cv-6728 (JMF), 2017 WL 1743853, at *1 (S.D.N.Y. May 4, 2017) (finding that plaintiffs' "motion for leave to amend amounts to a motion for reconsideration"). But even a motion for reconsideration "is not a vehicle for relitigating old issues, securing a rehearing on the merits, or otherwise taking a second bite at the apple." *Gustavia Home, LLC v. Rice*, 724 F. App'x 87, 88 (2d Cir. 2018). And Plaintiffs have made no effort to meet the "strict" standard for reconsideration—they have not identified "an intervening change of controlling law, the availability of new evidence, or the need to correct a clear error or prevent manifest justice." *U.S. Bank Nat. Assoc. v. Triaxx Asset Mgmt. LLC*, 352 F. Supp. 3d 242, 246 (S.D.N.Y. 2019). Plaintiffs may disagree with the Court's legal conclusions, but their alleged injury is still "the mere removal of identifying information from a copyrighted work—absent dissemination." Ord. at 7. OpenAI should therefore not be required "to battle for [the court's decision] again." *Off. Comm. Of Unsecured Creditors of Color Tile, Inc. v. Coopers & Lybrand, LLP*, 322 F.3d 147, 167 (2d Cir. 2003).

> **c.   Plaintiffs' copyright infringement analogy continues to fail.**

Even if the Court were to reevaluate Plaintiffs' copyright infringement analogy, it again fails on the merits because the proposed amendments still do not demonstrate the "necessary close historical or common-law analogue to Plaintiffs' alleged injury." *See* Ord. at 6.

8

As an initial matter, the analogy fails because the proposed amendments are about OpenAI's conduct, not about the injury Plaintiffs allegedly suffered as a result of that conduct. In this respect, Plaintiffs misunderstand the concreteness inquiry under *TransUnion*. There, the Supreme Court held that "the alleged *injury to the plaintiff*" must have a "close relationship to a harm traditionally recognized as providing a basis for a lawsuit in American courts." *TransUnion*, 594 U.S. at 424 (emphasis added). But Plaintiffs do not propose any amendments regarding the harm they suffered. Instead, Plaintiffs' proposed amendments consist almost entirely of new allegations that OpenAI's technical processes constitute copyright infringement.[1] Plaintiffs, however, still do not plead any registered copyright or assert a copyright infringement claim. Nor do they allege "any *actual* adverse effects stemming from" these technical processes. *See* Ord. at 7 (emphasis in original).

Instead, Plaintiffs mistakenly assume that there must be standing whenever there is a close relationship between a defendant's alleged conduct and a common law cause of action like copyright infringement. But the concreteness inquiry asks whether *Plaintiffs' asserted harm* "has a close relationship to a harm traditionally recognized as providing a basis for a lawsuit in American courts," not whether OpenAI's alleged conduct does. *TransUnion*, 594 U.S. at 417 (emphasis added). Plaintiffs argue that "even if DMCA violations are not always analogous to copyright infringement, they are on the particular facts of this case." Mot. at 4. But it is irrelevant

---

[1] *See, e.g.*, Mot. at 4 ("Plaintiffs will allege that Defendants removed Plaintiffs' CMI through a technical process that constitutes actual *prima facie* copyright infringement."); *id.* ("Defendants downloaded Plaintiffs' copyright-protected works and applied computer algorithms to create further copies with CMI deliberately removed."); *id.* ("Defendants cannot reasonably dispute that such copying is *prima facie* copyright infringement."); *id.* (arguing that standing for dissemination-less DMCA claims must exist "on the particular facts of this case, where Defendants committed actual *prima facie* copyright infringement in the course of violating the DMCA").

that the alleged conduct may share certain characteristics with conduct constituting copyright infringement. What matters under *TransUnion* is whether the "*actual* adverse effects stemming from this alleged DMCA violation" are analogous to the *actual* adverse effects stemming from copyright infringement. *See* Ord. at 7 (emphasis in original). Here, the Court has answered that question in the negative, at least absent allegations of dissemination of Plaintiff' own works. Ord. at 7. Accordingly, Plaintiffs proposed amendments are futile.

Further, Plaintiffs' proposed amendments do not address the defect that doomed their copyright infringement analogy the first time around, nor could they. The Court already decided that Section 1202 does not "grant the copyright owner the sole prerogative to decide how future iterations of the work may differ from the version the owner published," so it does not matter that Plaintiffs now allege new details about how, in their view, OpenAI interfered with that prerogative. *See* Ord. at 6. The point is that such interference is not the province of Section 1202(b). "Whether or not that type of injury satisfies the injury-in-fact requirement" in the abstract, "it is not the type of harm that has been 'elevated' by Section 1202(b)(i) of the DMCA." Ord. at 9. Accordingly, Plaintiffs' proposed amendments are futile regarding their copyright infringement analogy.

d.   Plaintiffs' new analogy to unjust enrichment is unavailing.

Rather than attempt to allege dissemination of their articles or articulate a different factual theory of how they were harmed, Plaintiffs instead draw a new legal analogy between its asserted DMCA injury and unjust enrichment. But this analogy, too, fails. As with their copyright infringement analogy, Plaintiffs' unjust enrichment argument is again "akin to that of the dissent in *TransUnion*." Ord. at 7. There, the dissent argued that the plaintiffs had standing because the statutory damages provision at issue had "a similar flavor" to unjust enrichment:

> TransUnion violated consumers' rights in order to create and sell a product to its clients. Reckless handling of consumer information and bungled responses to requests for information served a means to an end. And the end was financial gain

. . . Yet thanks to this Court, it may well be in a position to keep much of its ill-gotten gains.

TransUnion, 594 U.S. at 459 (Thomas, J., dissenting).

Plaintiffs' argument tracks this reasoning: "[OpenAI] violated [Plaintiffs'] rights in order to create and sell a product to its clients. Reckless [CMI removal] served a means to an end. And the end was financial gain." *Id.* Yet, this was not enough for the majority in *TransUnion*, and it is not enough here. *See TransUnion*, 594 U.S. at 417 ("No concrete harm, no standing.").

The Supreme Court also implicitly rejected Plaintiffs' theory in *Thole v. U.S. Bank, N.A.*, 590 U.S. 538 (2020). There, beneficiaries of a defined-contribution retirement plan sued under the Employee Retirement Income Security Act of 1974 (ERISA) for alleged mismanagement of the plan. 590 U.S. at 541. Crucially, the plaintiffs had so far received their monthly pension benefits, and they were "legally and contractually entitled to receive those same monthly payments for the rest of their lives." *Id.* at 540. The Court explained that if plaintiffs "were to win this lawsuit, they would still receive the exact same monthly benefits that they are already slated to receive, not a penny more." *Id.* at 541. Thus, the Court held that "the plaintiffs themselves have no concrete stake in the lawsuit," and, accordingly, "they lack Article III standing." *Id.* at 542.

The dissent, however, took the view that plaintiffs had standing under an unjust enrichment theory: a "beneficiary may sue a trustee for restitution or disgorgement, remedies that recognize the relevant harm as the trustee's wrongful gain." *Id.* at 558 (Sotomayor, J., dissenting). But this was not enough for the majority. *See id.* at 544 ("[T]he cause of action does not affect the Article III standing analysis."). Here, once again, Plaintiffs' argument is nearly a carbon copy of the argument raised in dissent; they even cite the same secondary source for support. In *Thole*, the dissent argued that "trustees are 'subject to liability' if they are unjustly enriched by 'a violation

11

of [a beneficiary]'s legally protected rights.'" *Id.* (quoting Restatement (Third) of Restitution and unjust Enrichment § 1 cmt. a.). Compare that to Plaintiffs' argument here:

> According to the applicable Restatement, "[a] person who is unjustly enriched at the expense of another is subject to liability." Restatement (Third) of Restitution and Unjust Enrichment, § 1. Critically, "at the expense of another" does not require an "observable loss," but can instead involve the "violation of the other's legally protected rights, without the need to show that the claimant has suffered a loss." *Id.* § 1 cmt. a.

Mot. at 6-7. There is no daylight between the argument raised in dissent in *Thole*, and Plaintiffs' unavailing argument here.

More generally, Plaintiffs' unjust enrichment analogy is wrong because it misses the distinction between the elements of unjust enrichment—which are defined as a matter of state law—and the "irreducible constitutional minimum of standing"—which is a matter of federal constitutional law. *Lujan v. Defenders of Wildlife*, 504 U.S. 555, 560 (1992). Plaintiffs may or may not be right that the state common law of unjust enrichment does not require an "observable loss" as one of its elements. Mot. at 7. But that does not change the fact that, as a matter of federal standing law, a "plaintiff's claim of injury in fact cannot be based solely on a defendant's gain; it must be based on a plaintiff's loss." *Silha v. ACT, Inc.*, 807 F.3d 169, 174-75 (7th Cir. 2015); *see also McNamara v. City of Chicago*, 138 F.3d 1219, 1221 (7th Cir. 1998) ("A plaintiff who would have been no better off had the defendant refrained from the unlawful acts of which the plaintiff is complaining does not have standing under Article III of the Constitution to challenge those acts in a suit in federal court.").

Accordingly, Plaintiffs lack Article III standing even if the alleged statutory violation "might [have] give[n] rise to liability in a lawsuit brought under the unjust enrichment cause of action." *Baehr v. Creig Northrop Team, P.C.*, 953 F.3d 244, 257 (4th Cir. 2020). This is because "concluding that a defendant's unjust enrichment always works a concrete injury to the plaintiff

in an action for statutory damages runs counter to *Spokeo's* mandate that 'a bare procedural

violation, divorced from any concrete harm' cannot 'satisfy the injury-in-fact requirement of

Article III.'"   *Id.* at 258 (quoting *Spokeo, Inc. v. Robins*, 578 U.S. 330, 341 (2016); *accord*

TransUnion, 594 U.S. at 440).

Consider another legal context that involves claims of mishandling a plaintiff's

information: data privacy cases. There, too, courts have regularly rejected theories of harm similar

to Plaintiffs'. For example, in *In re Google, Inc. Privacy Litig.*, plaintiffs claimed, *inter alia*, that

Google violated users' privacy rights by taking information from a user's Gmail account and using

it for other purposes, such as to personalize Google search results. No. C-12-01382-PSG, 2013

WL 6248499, at *1 (N.D. Cal. Dec. 3, 2013). Plaintiffs in that case argued that they were injured

because "Google did not compensate them for the substantial economic value of the []

information." *Id.* at *5. But the Court rejected this argument. It explained that "an allegation that

Google profited [from plaintiffs' information] is not enough equivalent to an allegation that such

profiteering deprived Plaintiffs of economic value from that same information." *Id.* So too, here.

Plaintiffs in this case "must do more than point to the dollars in a defendant's pocket; [they] must

sufficiently allege that in the process [they] lost dollars of their] own." *See id.*[2] They identify none,

and so their new analogy does not aid them in establishing standing.

---

[2] *See also Del Veccio v. Amazon.com, Inc.*, No. C11-366RSL, 2012 WL 1997697, at *9 (W.D.
Wash. June 1, 2012) ("The crux of an unjust enrichment claim is that a person who is unjustly
enriched *at the expense of another* is liable in restitution to the other. In regard to expense, . . .
Plaintiffs have failed to plausibly allege any monetary detriment.") (emphasis in original); *Low v.
LinkedIn Corp.*, No. 11-CV-01468-LHK, 2011 WL 5509848, at *4-5 (N.D. Cal. Nov. 11, 2011)
("Low alleges that his browsing history is personal property with market value and that he has
relinquished this valuable property without compensation to which he was due," but "has failed
to allege facts that demonstrate that he was economically harmed by LinkedIn's practices.");
LaCourt v. Specific Media, Inc., No. SACV 10-12560-GW, 2011 WL 1661532, at *4-5 (C.D.
Cal. Apr. 28, 2011) (in case about using cookies to track users' online activity without consent,
rejecting standing theory that "by taking and retaining Plaintiffs' personal information,

3. Plaintiffs' proposed amendments do not state a claim upon which relief can be granted.

Even if the Court finds that Plaintiffs have Article III standing, Plaintiffs' amendments are futile because they fail to remedy the other legal flaws in their proposed amended complaint. Plaintiffs once again fail to establish that they are "person[s] injured by a violation of section . . . 1202," and therefore fall outside the class of plaintiffs authorized to "bring a civil action" for such a violation. 17 U.S.C. § 1203. Further, Plaintiffs' proposed amendments are futile because they fail to plausibly allege that OpenAI acted with the requisite state of mind.

a. Plaintiffs still are not among those authorized to sue under 17 U.S.C. § 1203(a).

Plaintiffs are not "persons injured" within the meaning of the DMCA. Contrary to Plaintiffs' motion, the "question of constitutional standing is not the same as establishing [] standing to sue . . . as a 'person injured' under the DMCA." *CoxCom, Inc. v. Chaffee*, 536 F.3d 101, 107 n.7 (2d Cir. 2008). Accordingly, even if the Court finds that OpenAI's alleged unjust enrichment is sufficient for Article III purposes, Plaintiffs must *also* allege that they have been "injured by" the alleged violation of Section 1202(b). 17 U.S.C. § 1203 (allowing persons "injured by" an alleged violation of Section 1202 to sue). Plaintiffs' proposed amendments do not do so: the only new allegation of purported "injury" in the Proposed First Amended Complaint is that *OpenAI* unjustly benefited because the alleged CMI removal "involves fewer computational resources and therefore is far less expensive than retraining." PFAC ¶ 95. Because Plaintiffs still

---

Defendant has deprived Plaintiffs of this information's economic value"); *Goodman v. HTC America, Inc.*, No. C11-1793MJP, 2012 WL 2412070, at *7 (W.D. Wash. June 26, 2012) (regarding misappropriation of personally identifying information, finding that plaintiffs' allegation of harm "is overly vague because it relies on the abstract concept of opportunity costs, and Plaintiffs do not explain how they were deprived of the economic value of their personal information simply because their unspecified personal information was purportedly collected by a third party").

"have not alleged any actual adverse effects stemming from this alleged DMCA violation," their proposed amendments are futile. Ord. at 7; *see also Steele v. Bongiovi*, 784 F. Supp. 2d 94, 98 (D. Mass. 2011) (finding that plaintiff "cannot, as a matter of law, prevail on his DMCA claim" because "he was not injured by the alleged acts").

b. Plaintiffs' claim fails because they still have not adequately pled scienter.

Plaintiffs have not plausibly alleged that OpenAI intentionally removed CMI with the requisite state of mind. Section 1202(b)(1) contains two discrete scienter requirements: (1) that the defendant "intentionally remove[d]" CMI from copyrighted works, and (2) the defendant undertook such removal "knowing, or … having reasonable grounds to know[] that [the removal] will induce, enable, facilitate, or conceal an infringement [of copyright]." 17 U.S.C. § 1202(b); *see also* OpenAI MTD at 14-15. The "reasonable grounds to know[]" requirement demands "some identifiable connection between the defendant's actions and the infringement or the likelihood of infringement." *Victor Elias Photography, LLC v. Ice Portal, Inc.*, 43 F.4th 1313, 1325 (11th Cir. 2022). Plaintiffs attempt to satisfy the "identifiable connection" aspect of the "reasonable grounds to know" requirement by proffering five theories of liability. None withstands scrutiny.

i. CMI removal could not "conceal" non-public datasets.

The alleged absence of CMI from a non-public dataset cannot conceal anything. Plaintiffs' first theory is that OpenAI's alleged removal of CMI from their articles contained in internal training datasets might have concealed OpenAI's "use of copyrighted material in [its] training sets." PFAC ¶¶ 87–88. The problem for Plaintiffs, however, is that—as their own allegations establish—OpenAI does not "publish[] the contents of the training sets used to train any version of ChatGPT." *Id.* ¶ 34. That is presumably why Plaintiffs do not even attempt to explain how the alleged absence of CMI in an internal, non-public repository could conceivably "conceal[]"

15

anything.   Indeed, the only purpose of CMI is to "inform *the public* that something is copyrighted"—the absence of CMI in a non-public dataset cannot have any effect on what the public does or does not know. *Roberts v. BroadwayHD LLC*, 518 F. Supp. 3d 719, 737 (S.D.N.Y. 2021) (emphasis added). This is precisely why Judge Martinez-Olguin of the Northern District of California, when presented with the same question in a nearly identical context, dismissed a Section 1202(b)(1) claim for failure to establish a causal connection between the alleged CMI removal from an internal dataset and some "infringement." *Tremblay v. OpenAI, Inc.,* 716 F.Supp.3d 772, 778 (N.D. Cal. 2024).[3] Plaintiffs' proposed amendments are accordingly futile.

ii.   CMI removal could not "conceal" infringement in outputs.

Plaintiffs' second theory is that the removal of CMI during the training process "would result in ChatGPT providing responses to ChatGPT users that incorporated or regurgitated material . . . without revealing that those works were subject to Plaintiffs' copyrights." PFAC ¶ 89. Plaintiffs, in other words, suggest that OpenAI removed CMI from their articles allegedly included in OpenAI's training datasets because, at the time OpenAI created the datasets, OpenAI (1) knew that ChatGPT would regurgitate text from Plaintiffs' articles; (2) knew that including Plaintiffs' CMI in its training data would cause ChatGPT to include that CMI in any outputs that included material regurgitated from Plaintiffs' articles; (3) removed CMI from the training data in order to prevent the inclusion of that CMI in those outputs; and (4) did so in order to "conceal" the alleged infringement that results from such regurgitation. This theory fails for two independent reasons.

First, Plaintiffs propose amendments suggesting that OpenAI was aware of the *general phenomenon* of regurgitation—*i.e.*, LLMs outputting portions of training data it has

---

[3] After Judge Martinez-Olguin's ruling, the *Tremblay* plaintiffs dropped their DMCA claims. *See* First Consolidated Amended Complaint in *In re OpenAI ChatGPT Litig.*, No. 3:23-cv-03223, Dkt. 120 (N.D. Cal. Mar. 13, 2024).

"memorize[d]." PFAC ¶ 67 n.10 (relying on "OpenAI and Journalism" blog post). But nothing about OpenAI's alleged understanding of the general phenomenon of regurgitation says anything about OpenAI "having reasonable grounds to know" that its alleged removal of CMI from Plaintiffs' articles would "induce enable, facilitate, or conceal" infringement.

Second, Plaintiffs fail to allege facts to support two necessary premises to their contention that OpenAI had the requisite "reasonable grounds to know": (1) that including CMI in training datasets would cause ChatGPT to include that CMI in outputs, and (2) that excluding CMI from training datasets would have any effect on the contents of those outputs. Nowhere do Plaintiffs allege that, but for ChatGPT's alleged omission of Plaintiffs' CMI in training datasets, such CMI would have been included in response to the same instruction. Accordingly, Plaintiffs have failed to plausibly allege that OpenAI had "reasonable grounds to know" that its alleged exclusion of CMI from Plaintiffs' articles in training datasets had any "identifiable connection" to the inclusion or exclusion of CMI in potential ChatGPT outputs. *Victor Elias*, 43 F.4th at 1325.

iii.   CMI removal could not induce, enable, or facilitate users to infringe.

Plaintiffs' third theory is that "removing CMI would induce, enable, or facilitate users to infringe Plaintiffs' copyright by distributing infringing responses to a future audience." Mot. at 16. This theory is unavailing for at least the same reasons as their "concealment" theory: Plaintiffs do not allege that ChatGPT has ever disseminated one of Plaintiffs' articles, let alone that OpenAI had "reasonable grounds to know" that ChatGPT would do so. Nor do they allege that including CMI in training datasets would cause ChatGPT to include that CMI in outputs or that excluding CMI from training datasets would have any effect on the contents of those outputs.

Plaintiffs also fail to plausibly allege that the absence of CMI would "induce" users to distribute outputs that they otherwise would not have distributed. Plaintiffs rely on the wholly

conclusory allegation that "users of ChatGPT would be less likely to distribute ChatGPT responses if they were made aware of" CMI because those users "respect the copyrights of others or fear liability for copyright infringement." PFAC ¶ 92. But Plaintiffs offer no facts in support of this assumption, nor any facts tending to show that OpenAI had "reasonable grounds to know" that users behave in this way. Plaintiffs' allegation is thus an "unwarranted[] deduction of fact" and should be disregarded. *See First Nationwide Bank v. Gelt Funding Corp.*, 27 F.3d 763, 771 (2d Cir. 1994).

<div align="center">

iv.   CMI removal could not facilitate "training-based infringement."

</div>

Plaintiffs' fourth theory is that OpenAI removed CMI during the training process to "facilitate" the creation of a system that "replicate[s] how ordinary English speakers express themselves." PFAC ¶ 95 (speculating that, "[h]ad ChatGPT . . . been trained on [works with CMI], [it] would have falsely learned that ordinary English speakers convey [CMI] in situations when they do not"). But nothing in the Proposed First Amended Complaint suggests that the creation of a system that "replicate[s] how ordinary English speakers express themselves" is itself an "infringement" of copyright. *Id.*; 17 U.S.C. § 1202(b). Even if the removal of CMI "facilitate[d]" the creation of such a system, *see* PFAC ¶ 95, it would not necessarily follow that the same act of CMI removal "facilitate[d]" the acts of copyright infringement alleged in the complaint: *i.e.*, (1) the "downloading" of Plaintiffs' works, *id.* ¶ 62; (2) the "encoding" of those works "in computer memory," *id.* ¶ 33; (3) the "regurgitat[ion]" of those works in outputs, *id.* ¶ 67; or (4) the creation of supposedly unlawful "abridge[ments]" of those works, *id.* ¶ 68. There is, in other words, no reason to suspect that the removal of CMI during the training process made any of these alleged infringements any "easier" to carry out "undetected." *Stevens v. CoreLogic, Inc.*, 899 F.3d 666,

<div align="center">

18

</div>

676. Absent such an "identifiable connection," Plaintiffs' claim fails. *Victor Elias*, 43 F.4th at 1325.

> ### v. CMI removal could not facilitate "Defendants' copying."

Plaintiffs' proposed amendments do not plausibly allege that CMI removal facilitates the other purported acts of "copying" alleged in the complaint. Plaintiffs' final theory is that removing CMI facilitated OpenAI's alleged copying of Plaintiffs' works because "CMI-less works take up less data, so copying them requires fewer computational and storage resources." Mot. at 17. But this theory relies on an "unwarranted deduction[] of fact," *First Nationwide Bank*, 27 F.3d at 771: Plaintiffs have not explained why OpenAI's alleged copying is made any easier by saving the exceedingly miniscule "computational and storage resources" that correspond to author, title, and copyright notice information. As this Court has already recognized, Plaintiffs allege that "ChatGPT has been trained on a scrape of most of the internet, which includes massive amounts of information from innumerable sources on almost any given subject." Ord. at 8. Plaintiffs fail to allege why author, title, and copyright notice information would make any appreciable difference among the "massive amounts of information" alleged to constitute OpenAI's training data. *Id.*

In any event, this theory directly contradicts Plaintiffs' other allegations that OpenAI "first download[s] and save[s] the relevant webpage before extracting data from it" because OpenAI allegedly "anticipate[s] a possible future need to regenerate the dataset (e.g., if the dataset becomes corrupted), and it is cheaper to save a copy than it is to recrawl all the data." PFAC ¶ 46. Plaintiffs cannot plausibly allege that CMI removal "makes it easier to copy Plaintiffs' works," Mot. at 17, while simultaneously alleging that OpenAI copies Plaintiffs' entire works, with CMI included.

B.   **Plaintiffs are not entitled to jurisdictional discovery.**

Plaintiffs' failure to allege facts about their *own* harm does not warrant jurisdictional discovery. Plaintiffs seek discovery to show that their alleged works have been disseminated by OpenAI. *See* Mot. at 11. But a plaintiff is not entitled to jurisdictional discovery if it cannot show that the requested discovery is "likely to produce the facts needed to withstand a Rule 12(b)(1) motion." *Molchatsky v. United States*, 778 F. Supp. 2d 421, 438 (S.D.N.Y. 2011), *aff'd*, 713 F.3d 159 (2d Cir. 2013).

Plaintiffs have not made that showing because their assertions are speculative. This Court has already made the common-sense observation that "[g]iven the quantity of information contained in" OpenAI's training data, "the likelihood that ChatGPT would output plagiarized content from one of Plaintiffs' articles seems remote." Ord. at 9. Plaintiffs' proposed amendments do not allege any facts to contradict this "remote[ness]". Plaintiffs admit that, despite their best efforts, they have failed to cause ChatGPT to output their works at all. PFAC ¶ 67.

Plaintiffs argue that it "is not feasible for Plaintiffs to identify which articles were disseminated" absent discovery. Mot. at 11. But this conclusion does not follow from Plaintiffs' own allegations. Indeed, their amendments confirm that several *other* plaintiffs have been able to allege "[e]xamples of such regurgitations" and "abridgements" without the benefit of jurisdictional discovery. *See Daily News LP v. Microsoft Corp.*, No. 24-cv-03285 (S.D.N.Y. Apr. 30, 2024), and *The Center for Investigative Reporting, Inc. v. OpenAI, Inc.*, No. 24-cv-4872 (S.D.N.Y. Sept. 9, 2024). This is not to say that the examples in those other cases meet the mark—but only to say that in this case, they have shown the Court even less.

In any event, Plaintiffs' only support for jurisdictional discovery is that their articles appear in OpenAI's training data. Mot. at 11. But Plaintiffs allege that OpenAI's models are trained on "a scrape of most of the internet," PFAC ¶ 53, "which includes massive amounts of information

20

from innumerable sources on almost any given subject," Ord. at 8. Accordingly, viewing Plaintiffs' allegations in the light most favorable to them, the mere appearance of Plaintiffs' works in OpenAI's training data does not justify jurisdictional discovery, especially considering Plaintiffs' failure to allege examples of dissemination of its articles. *See Amidax Trading Group v. S.W.I.F.T. SCRL*, 607 F. Supp. 2d 500, 509 n.5 (S.D.N.Y. 2009) ("Given the insufficiency of plaintiff's allegations of standing, the Court finds that jurisdictional discovery is not warranted.").

Plaintiffs argue that they need discovery to determine whether they were injured. But Plaintiffs' inability to recognize their own injury without discovery only confirms that their purported harm is speculative. *Cf. Transunion*, 594 U.S. at 438 (finding the fact that a plaintiff would not realize they had been "injured" for purposes of Article III until "they received a check compensating them for their supposed 'injury'" supported the argument that any alleged harm was speculative). Plaintiffs' request for jurisdictional discovery should be denied.

V.   CONCLUSION

For the foregoing reasons, the Court should deny Plaintiffs' motion for leave to file an amended complaint, deny Plaintiffs' belated request for jurisdictional discovery, and dismiss the case with prejudice.

Dated: December 20, 2024

MORRISON & FOERSTER LLP

By: */s/ Joseph C. Gratz*
    Joseph C. Gratz (*pro hac vice*)
    JGratz@mofo.com
    Andrew L. Perito (*pro hac vice*)
    APerito@mofo.com
    Vera Ranieri (*pro hac vice*)
    VRanieri@mofo.com
    425 Market Street
    San Francisco, CA 94105-2482
    Telephone: 415.268.7000
    Facsimile: 415.268.7522

    Rose S. Lee (*pro hac vice*)
    Roselee@mofo.com
    707 Wilshire Boulevard, Suite 6000
    Los Angeles, CA 90017-3543
    Telephone: 213.892.5200
    Facsimile: 213.892.5454

    Carolyn M. Homer (*pro hac vice*)
    Cmhomer@mofo.com
    2100 L Street NW
    Washington, D.C. 20037
    Telephone: 202.650.1500
    Facsimile: 202.887.0763

    Emily C. Wood
    EWood@mofo.com
    Eric K. Nikolaides
    ENikolaides@mofo.com
    250 W. 55th Street
    New York, NY 10019-9601
    Telephone: 212.468.8000
    Facsimile: 212.468.7900

    Attorneys for Defendants
    OPENAI, INC., OPENAI GP, LLC,
    OPENAI, LLC, OPENAI OPCO LLC,
    OPENAI GLOBAL LLC, OAI
    CORPORATION, LLC, and OPENAI
    HOLDINGS, LLC

Dated: December 20, 2024   LATHAM & WATKINS LLP

By: */s/ Elana Nightingale Dawson*
    Andrew Gass (*pro hac vice*)
    andrew.gass@lw.com
    Joseph R. Wetzel (*pro hac vice*)
    joe.wetzel@lw.com
    505 Montgomery Street, Suite 2000
    San Francisco, CA 94111
    Telephone: 415.391.0600
    Facsimile: 415.395.8095

    Sarang Damle
    sy.damle@lw.com
    Elana Nightingale Dawson (*pro hac vice*)
    elana.nightingaledawson@lw.com
    555 Eleventh Street NW
    Suite 100
    Washington, DC 20004
    Telephone: 202.637.2200
    Facsimile: 202.637.2201

    Allison Levine Stillman
    alli.stillman@lw.com
    Rachel R. Blitzer
    rachel.blitzer@lw.com
    Herman H. Yue
    herman.yue@lw.com
    Michael A. David
    michael.david@lw.com
    Luke Budiardjo
    luke.budiardjo@lw.com
    1271 Avenue of the Americas
    New York, NY 10020
    Telephone: 212.906.1200
    Facsimile: 212.751.4864

    Attorneys for Defendants
    OPENAI, INC., OPENAI GP, LLC,
    OPENAI, LLC, OPENAI OPCO LLC,
    OPENAI GLOBAL LLC, OAI
    CORPORATION, LLC, and OPENAI
    HOLDINGS, LLC

Dated: December 20, 2024   KEKER, VAN NEST & PETERS LLP

By: */s/ Thomas E. Gorman*
    Robert A. Van Nest (*pro hac vice*)
    RVanNest@keker.com
    R. James Slaughter (*pro hac vice*)
    RSlaughter@keker.com
    Paven Malhotra
    PMalhotra@keker.com
    Michelle S. Ybarra (*pro hac vice*)
    MYbarra@keker.com
    Nicholas S. Goldberg (*pro hac vice*)
    NGoldberg@keker.com
    Thomas E. Gorman (*pro hac vice*)
    TGorman@keker.com
    Katie Lynn Joyce (*pro hac vice*)
    KJoyce@keker.com
    Christopher S. Sun (*pro hac vice*)
    CSun@keker.com
    Andrew S. Bruns (*pro hac vice*)
    abruns@keker.com
    Andrew Dawson (*pro hac vice*)
    adawson@keker.com
    Edward A. Bayley
    ebayley@keker.com
    633 Battery Street
    San Francisco, California 94111-1809
    Telephone: 415.391.5400
    Facsimile: 415.397.7188

    Attorneys for Defendants
    OPENAI, INC., OPENAI GP, LLC,
    OPENAI, LLC, OPENAI OPCO LLC,
    OPENAI GLOBAL LLC, OAI
    CORPORATION, LLC, and OPENAI
    HOLDINGS, LLC