Co. v. Raimondo, 40 F.4th 716, 722 (D.C. Cir. 2022) (citation omitted).

[61] The defendants argue that the ultra vires claim should be dismissed because it is "coextensive" with the complaint's Privacy Act claims and the plaintiffs have failed to plead that Exception (b)(1) of the Privacy Act did not permit the OPM Defendants to disclose OPM records to the DOGE Defendants. This argument fails. The complaint alleges a massive disclosure of the OPM records of tens of millions of Americans to unvetted and untrained individuals who had no legal right to access those records, in wholesale disregard of the Privacy Act. It pleads that this intrusion was directed and controlled by the DOGE Defendants, including individuals in USDS or associated with USDS. The DOGE Defendants have no statutory authority with respect to OPM records, such that these alleged actions were "blatantly lawless." Fed. Express Corp., 39 F.4th at 764 (citation omitted). The complaint adequately pleads that the DOGE Defendants "plainly and openly crossed a congressionally drawn line in the sand." Id. at 765.

Conclusion

The defendants' March 14, 2025 motion to dismiss under Rule 12(b)(6) is granted as to the complaint's First and Second Claims, except insofar as they are a predicate to the complaint's other claims. The defendants' motion to dismiss under Rule 12(b)(1) is denied. The defendants' motion to dismiss claims Three, Four, and Five under Rule 12(b)(6) is denied.

The NEW YORK TIMES COMPANY, Plaintiff,

v.

MICROSOFT CORPORATION, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC, Defendants.

Daily News LP, et al., Plaintiff,

v.

Microsoft Corporation, OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC, Defendants.

The Center for Investigative Reporting, Inc., Plaintiff,

v.

OpenAI, Inc., OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC, OpenAI Holdings, LLC, and Microsoft Corporation, Defendants.

23-cv-11195 (SHS)
24-cv-3285 (SHS)
24-cv-4872 (SHS)

United States District Court, S.D. New York.

Signed April 4, 2025

Background: First news organization filed amended complaint against technology company and artificial intelligence (AI) company asserting claims for direct copyright infringement, contributory copyright infringement, and violation of the Digital Millennium Copyright Act (DMCA), and second and third news organizations each filed separate complaints asserting claims for direct copyright infringement, contrib-

utory copyright infringement, vicarious copyright infringement, violation of the DMCA, common law unfair competition by misappropriation, and trademark dilution, and seeking monetary and injunctive relief. Defendants filed motions to dismiss for failure to state a claim, lack of standing, and on statute of limitations grounds.

Holdings: The District Court, Sidney H. Stein, J., held that:

(1) AI company failed to demonstrate that organizations discovered, or should have discovered, company's alleged infringement, as would support applying discovery rule to bar claims for direct copyright infringement;

(2) organizations plausibly alleged that end-user infringement of their copyrighted material had taken place, as required for claims for contributory copyright infringement;

(3) organizations plausibly alleged that harm incurred as result of alleged violations of the DMCA were sufficiently concrete to satisfy injury-in-fact requirement for Article III standing;

(4) second organization failed to plausibly allege that copyright management information (CMI) was removed from its articles, as required for claim for violation of the DMCA;

(5) allegations in complaints were insufficient to state claims against technology company for violation of the DMCA;

(6) second and third organizations failed to plausibly allege that companies were "free-riding" on organizations' content, as required for misappropriation claims to survive preemption by the Copyright Act under "hot news" exception; but

(7) organizations plausibly alleged that their publications were famous, as required for claims for trademark dilution under the Lanham Act.

Motions granted in part and denied in part.

1. Federal Civil Procedure ⚏1831

The choice between two plausible inferences that may be drawn from factual allegations is not a choice to be made by the district court on a motion to dismiss for failure to state a claim; fact-specific questions cannot be resolved on the pleadings. Fed. R. Civ. P. 12(b)(6).

2. Limitation of Actions ⚏95(7)

Under the discovery rule, a claim of copyright infringement does not accrue until the copyright holder discovers, or with due diligence should have discovered, the infringement. 17 U.S.C.A. § 507(b).

3. Limitation of Actions ⚏95(7)

In the context of an action for direct copyright infringement, the discovery rule does not impose on copyright holders a general duty to police the internet to uncover infringement. 17 U.S.C.A. §§ 501, 507(b).

4. Limitation of Actions ⚏95(7)

In the context of the discovery rule in an action for direct copyright infringement, a copyright holder's general diligence or allegations of diligence in seeking out and litigating infringements, alone, are insufficient to make it clear that the holder's particular claims in any given case should have been discovered more than three years before the action's commencement. 17 U.S.C.A. §§ 501, 507(b).

5. Limitation of Actions ⚏95(7)

To establish constructive notice of copyright infringement, for purposes of the discovery rule in an action for direct copyright infringement, the defendant must identify specific facts or circumstances that would have prompted such an inquiry by the copyright holder into the

alleged infringing activity. 17 U.S.C.A. §§ 501, 507(b).

6. Limitation of Actions ☞95(1)

There is no "sophisticated plaintiff" exception to the discovery rule.

7. Limitation of Actions ☞95(7)

In the context of the discovery rule in an action for direct copyright infringement, the date on which a copyright holder, with the exercise of due diligence, would have discovered an infringement, or whether the alleged date of discovery reflected a lack of due diligence, is a fact-intensive inquiry that cannot be determined from the general nature of a copyright holder's sophistication alone. 17 U.S.C.A. §§ 501, 507(b).

8. Federal Civil Procedure ☞1754

Dismissal of a copyright infringement claim on statute of limitations grounds at the pleadings stage is only appropriate when it is clear from the face of the complaint, and matters of which the court may take judicial notice, that the plaintiff's claims are barred as a matter of law; however, where there is even some doubt as to whether dismissal is warranted, a court should not grant a motion to dismiss for failure to state a claim on statute of limitations grounds. 17 U.S.C.A. §§ 501, 507(b); Fed. R. Civ. P. 12(b)(6).

9. Limitation of Actions ☞95(7)

Artificial intelligence (AI) company failed to demonstrate that news organizations discovered, or should have discovered, company's alleged infringement of their news content more than three years before organizations filed their complaints, as would support applying the discovery rule to bar claims for direct copyright infringement; while the complaints alleged company trained its large language models (LLM) on datasets that included organizations' works created more than three years

before the complaints were filed, and identified articles suggesting that it was common knowledge that the datasets included the works, company failed to explain why organizations would have known of the articles or why the articles were sufficient to put organizations on notice of the infringing conduct. 17 U.S.C.A. §§ 501, 507(b).

10. Copyrights and Intellectual Property ☞685

An individual or entity may be held liable as a contributory copyright infringer if that individual or entity, with knowledge of the infringing activity, induces, causes, or materially contributes to the infringing conduct of another.

11. Copyrights and Intellectual Property ☞685

To plausibly allege contributory copyright infringement under a theory of material contribution, a plaintiff must show (1) direct infringement by a third party, (2) that the defendant had knowledge of the infringing activity, (3) and that the defendant materially contributed to the third party's infringement.

12. Copyrights and Intellectual Property ☞686

The knowledge standard for a claim of contributory copyright infringement is an objective one; contributory infringement liability is imposed on persons who know or have reason to know of the direct infringement of a copyright.

13. Copyrights and Intellectual Property ☞686

More than a generalized knowledge by the defendant of the possibility of infringement is required to meet the knowledge requirement of a claim of contributory copyright infringement.

14. Copyrights and Intellectual Property
    ⊗686

Knowledge of specific infringements is not required to support a finding of contributory copyright infringement.

15. Copyrights and Intellectual Property
    ⊗687

To establish that the defendant materially contributed to a third party's infringement of a copyright, for purposes of a claim of contributory copyright infringement, the complaint must show that the defendant encouraged or assisted others' infringement or provided machinery or goods that facilitated infringement.

16. Copyrights and Intellectual Property
    ⊗972

News organizations plausibly alleged that end-user infringement of their copyrighted material had taken place, as required for claims against technology company and artificial intelligence (AI) company for contributory copyright infringement, which partnered to develop and train large language models (LLM) using organization's news content; organizations' complaints alleged that "widely publicized" instances of copyright infringement occurred after companies released their text-generating chatbots, and included more than 100 examples of infringing outputs from LLMs, all of which raised a reasonable expectation that discovery would reveal evidence of additional examples of third-party infringement.

17. Copyrights and Intellectual Property
    ⊗972

News organizations plausibly alleged that technology company and artificial intelligence (AI) company had actual or constructive knowledge of end-user infringement of their copyrighted content, as required for claims against companies for contributory copyright infringement, alleging companies used organizations' content to train their AI large language models (LLM); complaints alleged that companies knew they were using copyrighted works to train their LLMs and were fully aware of organizations' protected interests in their works, that first organization informed companies "that their tools infringed its copyrighted works," and companies knew that their use of the works "would inevitably result in the unauthorized display of such works" in response to third-party queries.

18. Federal Civil Procedure ⊗103.2,
    103.3

In order to establish Article III standing, the United States Constitution requires a plaintiff to show: (1) that he suffered an injury in fact that is concrete, particularized, and actual or imminent, (2) that the injury was likely caused by the defendant, and (3) that the injury would likely be redressed by judicial relief. U.S. Const. art. 3, § 2, cl. 1.

19. Federal Civil Procedure ⊗103.2

For an injury to be concrete, as required for Article III standing, the injury must be actual or imminent, not conjectural or hypothetical. U.S. Const. art. 3, § 2, cl. 1.

20. Federal Civil Procedure ⊗103.2

While Congress may elevate harms that exist in the real world before Congress recognized them to actionable legal status, it may not simply enact an injury into existence, for purposes of Article III standing, using its lawmaking power to transform something that is not remotely harmful into something that is. U.S. Const. art. 3, § 2, cl. 1.

21. Federal Civil Procedure ⊗103.2

Regardless of the existence vel non of a statutory cause of action, courts must independently decide whether a plaintiff

has suffered a concrete harm, for purposes of standing under Article III; to do so, courts look to 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, namely whether plaintiffs have identified a close historical or common-law analogue for their asserted injury. U.S. Const. art. 3, § 2, cl. 1.

22. Copyrights and Intellectual Property
$\Longleftrightarrow$**972**

News organizations plausibly alleged that the harm they incurred as a result of the alleged violations of the Digital Millennium Copyright Act (DMCA) by technology company and artificial intelligence (AI) company were sufficiently concrete to satisfy the injury-in-fact requirement for Article III standing, in action alleging companies used organizations' content to train their AI large language models (LLM); complaints alleged that companies' use of their content constituted copyright infringement, which was a harm traditionally recognized as providing a basis for lawsuits and provided a close historical analogue to the harm caused by a DMCA violation, as both claims were grounded in property rights and exercised Congress' power to "promote the Progress of Science and useful Arts." U.S. Const. art. 1, § 8, cl. 8; U.S. Const. art. 3, § 2, cl. 1; 17 U.S.C.A. §§ 1202(b)(1), 1202(b)(3).

23. Copyrights and Intellectual Property
$\Longleftrightarrow$**664**

Claims under the Digital Millennium Copyright Act (DMCA) protect against harms caused by unauthorized removal of copyright management information from copyrighted work; traditional copyright infringement claims protect against, among other things, unauthorized reproduction and distribution of protected works. 17 U.S.C.A. § 1202.

24. Copyrights and Intellectual Property
$\Longleftrightarrow$**945**

For purposes of Article III standing, copyright infringement, a harm traditionally recognized as providing a basis for lawsuits in American courts, provides an appropriate close historical or common-law analogue to the harm caused by a violation of the Digital Millennium Copyright Act (DMCA). U.S. Const. art. 3, § 2, cl. 1; 17 U.S.C.A. § 1202.

25. Federal Civil Procedure $\Longleftrightarrow$**103.3**

To satisfy the causation requirement for Article III standing, a plaintiff must establish that the plaintiff's injury likely was caused or likely will be caused by the defendant's conduct; put differently, there must be a causal connection between the injury and the conduct complained of, the injury has to be fairly traceable to the challenged action of the defendant, and not the result of the independent action of some third party not before the court. U.S. Const. art. 3, § 2, cl. 1.

26. Copyrights and Intellectual Property
$\Longleftrightarrow$**973**

News organizations plausibly alleged that the harm they incurred was fairly traceable to the alleged removal of the copyright management information (CMI) from organizations' copyrighted works by technology company and artificial intelligence (AI) company, as required to satisfy the injury-in-fact requirement for Article III standing, in action for violation of the Digital Millennium Copyright Act (DMCA), alleging companies used organizations' content to train their AI large language models (LLM); complaints alleged that companies deprived organizations of licensing and subscription revenue by allowing companies to directly provide the works to end users through regurgitating outputs, while concealing that companies infringed organizations' copyrights to

generate those outputs. U.S. Const. art. 3, § 2, cl. 1; 17 U.S.C.A. §§ 1202(b)(1), 1202(b)(3).

27. Copyrights and Intellectual Property
    ☞973

News organizations plausibly alleged that they had statutory standing to bring claims against technology company and artificial intelligence (AI) company for violation of the Digital Millennium Copyright Act (DMCA), alleging companies used organizations' content to train their AI large language models (LLM); complaints alleged that companies deprived organizations of licensing and subscription revenue by allowing companies to directly provide the works to end users through regurgitating outputs, while concealing that companies infringed organizations' copyrights to generate those outputs. 17 U.S.C.A. §§ 1202(b)(1), 1202(b)(3), 1203(a).

28. Copyrights and Intellectual Property
    ☞664

To establish a violation of the provision of the Digital Millennium Copyright Act's (DMCA) prohibiting intentionally removing or altering any copyright management information (CMI), a plaintiff must show (1) the existence of CMI on the allegedly infringed work, (2) the removal or alteration of that information, and (3) that the removal was intentional. 17 U.S.C.A. § 1202(b)(1).

29. Copyrights and Intellectual Property
    ☞973

Second news organization failed to plausibly allege that copyright management information (CMI) was removed from its articles by artificial intelligence (AI) company during the process of training company's large language models (LLM), as required for a claim for violation of the Digital Millennium Copyright Act (DMCA); complaint merely alleged that because certain of the LLM's regurgi-

tating outputs lacked CMI then CMI was removed, but, it was plausible that CMI remained on the articles but simply did not appear in the outputs, as the regurgitating outputs all contained excerpts of the articles, rather than complete or substantially complete copies, and the complaint did not include any specific detail on how CMI was allegedly removed during the training process. 17 U.S.C.A. § 1202(b)(1).

30. Copyrights and Intellectual Property
    ☞973

Third news organization plausibly alleged that copyright management information (CMI) was removed from its articles by artificial intelligence (AI) company during the process of training company's large language models (LLM), as required for a claim for violation of the Digital Millennium Copyright Act (DMCA); complaint alleged that AI company used two certain extractors in creating the training dataset, which intentionally removed the organization's CMI from its works scraped from its website, and explained that the first extractor removed copyright notices "as part of the process of extracting the text content of a website," and alleged that the second extractor separated and extracted news article text on the webpages while removing CMI. 17 U.S.C.A. § 1202(b)(1).

31. Copyrights and Intellectual Property
    ☞973

First news organization plausibly alleged that copyright management information (CMI) was removed from articles on its website by artificial intelligence (AI) company during the process of training company's large language models (LLM), as required for a claim for violation of the Digital Millennium Copyright Act (DMCA); complaint alleged that company, in developing a dataset to train the LLMs, used sets of extractors to extract text from websites, that the extractors were de-

signed to separate the main article content from other parts, including footers and copyright notices, that company allowed an extractor to make further copies only of the main article content, and that another extractor was incapable of extracting copyright notices and footers. 17 U.S.C.A. § 1202(b)(1).

32. Copyrights and Intellectual Property
✎973

Third news organization plausibly alleged that artificial intelligence (AI) company's removal of copyright management information (CMI) from the articles on organization's website during the process of training company's large language models (LLM) was intentional, as required for a claim for violation of the Digital Millennium Copyright Act (DMCA); complaint explained that company intentionally used certain content extractors that were designed to remove CMI from the works they scraped from the internet. 17 U.S.C.A. § 1202(b)(1).

33. Copyrights and Intellectual Property
✎973

First news organization plausibly alleged that artificial intelligence (AI) company's removal of copyright management information (CMI) from the articles on organization's website during the process of training company's large language models (LLM) was intentional, as required for a claim for violation of the Digital Millennium Copyright Act (DMCA); complaint alleged that because, by the time of its scraping, certain content extractors used by company were publicly known to remove author, title, copyright notices, and footers, and that AI company intentionally and knowingly removed the copyright management information while assembling a dataset to train the LLMs, and that AI company used highly skilled data scientists who would know how the extractors worked. 17 U.S.C.A. § 1202(b)(1).

34. Copyrights and Intellectual Property
✎973

First and third news organization plausibly alleged that artificial intelligence (AI) company knew or had reason to know that its removal of copyright management information (CMI) from the articles on organization's website during the process of training companies' large language models (LLM) would induce, enable, facilitate, or conceal copyright infringement, as required for a claim for violation of the Digital Millennium Copyright Act (DMCA); complaint alleged that AI company publicly acknowledged that it used copyrighted works to train its models, and that AI company was aware of the possibility of end-user infringement through the generating of regurgitating outputs, as it agreed to indemnify end users accused of infringement. 17 U.S.C.A. § 1202(b)(1).

35. Copyrights and Intellectual Property
✎973

Allegations in the complaints brought by news organizations were insufficient to state claims against technology company for violation of the provision of the Digital Millennium Copyright Act (DMCA) prohibiting intentionally removing or altering any copyright management information (CMI), alleging technology company used content scraped from organizations' websites to train their artificial intelligence (AI) large language models (LLM); complaints merely alleged that technology company partnered with AI company to train and develop LLMs, and that technology company provided the cloud computing system which AI company used to train its models, but did not allege that technology company removed any CMI from organizations' works. 17 U.S.C.A. § 1202(b)(1).

36. Copyrights and Intellectual Property
✎664

To establish a violation of the provision of the Digital Millennium Copyright

Act (DMCA) prohibiting the distribution of works or copies of works knowing that copyright management information (CMI) has been removed or altered without authority of the copyright owner, a plaintiff must show (1) the existence of CMI in connection with a copyrighted work, and (2) that a defendant distributed works or copies of works, (3) while knowing that CMI has been removed or altered without authority of the copyright owner or the law, and (4) while knowing, or having reasonable grounds to know that such distribution will induce, enable, facilitate, or conceal an infringement.   17 U.S.C.A. § 1202(b)(3).

37. Copyrights and Intellectual Property ☞663

Second and third news organizations failed to plausibly allege that technology company and artificial intelligence (AI) company distributed copies of organizations' works, as required for claims for violation of the Digital Millennium Copyright Act (DMCA); complaint merely alleged that companies distributed regurgitations from the AI large language models (LLM) they trained using content scraped from organizations' websites, however, the regurgitations included only excerpts of the underlying articles, the excerpts were often generated in response to multiple prompts, and the excerpts did not represent exact or substantial reproductions of the originals, as they outputted in a different order than they appeared in the original articles.   17 U.S.C.A. § 1202(b)(3).

38. Copyrights and Intellectual Property ☞973

First news organization failed to plausibly allege that technology company, which partnered with artificial intelligence (AI) company to train certain AI large language models (LLM) using articles scraped from organization's websites, distributed to AI company copies of organi-

zation's works without their copyright management information (CMI), as would support a claim against technology company for violation of the Digital Millennium Copyright Act (DMCA); complaint merely asserted that technology company's investment in AI company, its provision of "the data center and bespoke supercomputing infrastructure used to train" an LLM, and its statement that "we have the data," gave rise to the inference that companies distributed copies of organization's works with each other.   17 U.S.C.A. § 1202(b)(3).

39. Copyrights and Intellectual Property ☞973

First news organization failed to plausibly allege that artificial intelligence (AI) company, which partnered with technology company to train certain AI large language models (LLM) using articles scraped from organization's websites, distributed to technology company copies of organization's works without their copyright management information (CMI), as would support a claim against AI company for violation of the Digital Millennium Copyright Act (DMCA); complaint merely detailed the general business relationship between companies, but failed to indicate that AI company "distributed" "copies" of organization's articles to technology company, and nowhere in the complaint did organization allege when, why, or how AI company would have distributed copies of the works to technology company.   17 U.S.C.A. § 1202(b)(3).

40. Copyrights and Intellectual Property ☞214
Federal Preemption ☞63

Certain forms of commercial misappropriation otherwise within the general scope requirement of the Copyright Act will survive preemption when those claims include extra elements instead of, or in

addition to, the acts of reproduction, performance, distribution or display that lie within the general scope of copyright. 17 U.S.C.A. § 301.

41. Antitrust and Trade Regulation ☞14

    Federal Preemption ☞63

To make out a "hot news" misappropriation claim that survives preemption under the Copyright Act, a plaintiff must at a minimum allege the following extra elements: (1) the time-sensitive value of factual information, (2) the "free-riding" by a defendant, and (3) the threat to the very existence of the product or service provided by the plaintiff. 17 U.S.C.A. § 301.

42. Antitrust and Trade Regulation ☞14

    Federal Preemption ☞63

Second and third news organizations failed to plausibly allege that technology company and artificial intelligence (AI) company, which partnered together to train AI large language models (LLM) using organizations' news content, were "free-riding" on organizations' content, as required for organizations' claims for unfair competition by misappropriation under New York common law to survive preemption by the Copyright Act under the "hot news" exception; complaints did not allege that companies did not identify the source of the content, as the prompts to the LLM listed in the complaint asked the LLM to provide the text of a specific article and included the title of the article in the prompt itself, and many prompts also provided the LLM with the name of the article's publisher. 17 U.S.C.A. § 301.

43. Antitrust and Trade Regulation ☞14

    Federal Preemption ☞63

Second and third news organizations failed to plausibly allege that organiza-

tions' content that technology company and artificial intelligence (AI) company used to train AI large language models (LLM) was time sensitive, as required for organizations' claims of unfair competition by misappropriation under New York common law to survive preemption by the Copyright Act under the "hot news" exception; nowhere did organizations allege that the LLM's outputs regurgitated each article precisely at the point where the profit was to be reaped. 17 U.S.C.A. § 301.

44. Antitrust and Trade Regulation ☞14

    Federal Preemption ☞63

Second news organization failed to plausibly allege that reviews from organization's publication that technology company and artificial intelligence (AI) company used to train AI large language models (LLM) were time sensitive, as required for organizations' claims of unfair competition by misappropriation under New York common law to survive preemption by the Copyright Act under the "hot news" exception; complaint did not allege that the reviews were breaking news, but rather that they were recommendations resulting from journalists spending "tens of thousands of hours each year researching and testing products to ensure that they recommend only the best," and the portion of the complaint discussing the reviews was separate from the complaint's "Breaking News" section. 17 U.S.C.A. § 301.

45. Antitrust and Trade Regulation ☞14

    Federal Preemption ☞63

Second news organization failed to plausibly allege that technology company and artificial intelligence (AI) company, which partnered together to train AI large language models (LLM) using organization's news content, were "free-riding" on

organizations' content, as required for organizations' claims of unfair competition by misappropriation under New York common law to survive preemption by the Copyright Act under the "hot news" exception; references to the reviews listed in the complaint all referenced the publication by name and attributed their recommendations to that publication, and recommendations outputted by the LLMs would have little value without their attribution to the publication, as the prompts specifically asked the LLMs to provide recommendations from the publication. 17 U.S.C.A. § 301.

46. Trademarks ⚖1463, 1466

To state a claim for trademark dilution under the Lanham Act, a plaintiff must allege: (1) the mark is famous, (2) the defendant's use of the mark is made in commerce, (3) the defendant used the mark after the mark is famous, and (4) the defendant's use of the mark is likely to dilute the quality of the mark by blurring or tarnishment. Lanham Trade-Mark Act § 43, 15 U.S.C.A. § 1125(c).

47. Trademarks ⚖1585

In pleading federal trademark dilution, a complaint's spare, conclusory allegations that the trademark holder has expended substantial time, effort, money, and resources advertising and promoting a trademarked product does not suffice to plausibly allege that a trademark is famous. Lanham Trade-Mark Act § 43, 15 U.S.C.A. § 1125(c)(2)(A).

48. Trademarks ⚖1468, 1585

A single, conclusory allegation that a trademark is widely recognized by the general public is not sufficient to plead federal trademark dilution; rather, a complaint plausibly alleges that a trademark is famous when the allegations include attributes of the mark such as nationwide recognition and respect, continuous and pervasive use of the mark, substantial investments in promoting and advertising the mark throughout the United States and internationally, significant publicity relating to the marks, and that products bearing the plaintiff's marks are sold throughout the United States. Lanham Trade-Mark Act § 43, 15 U.S.C.A. § 1125(c)(2)(A).

49. Trademarks ⚖1468

News organizations plausibly alleged that their publications were famous, as required for claims for trademark dilution under the Lanham Act against technology company and artificial intelligence (AI) company that used publications' content to train their large language models (LLM); complaint alleged that the publication has been in circulation for more than 100 years, that several trademarks were federally registered, that organizations collectively owned over 40,000 copyright registrations for works published under the trademarks, that the publications were circulated throughout all 50 states, that their news stories were featured by national news outlets and had received significant attention and praise from news outlets, and that the publications had achieved international fame. Lanham Trade-Mark Act § 43, 15 U.S.C.A. § 1125(c)(2)(A).

50. Commerce ⚖56

The dormant Commerce Clause prevents a state from using its laws to discriminate purposefully against out-of-state economic interests. U.S. Const. art. 1, § 8, cl. 3.

51. Commerce ⚖55, 60(1)

For purposes of assessing whether the dormant Commerce Clause is violated, absent discrimination, a state may exclude from its territory, or prohibit the sale therein of any articles which, in its judgment, fairly exercised, are prejudicial to

the interests of its citizens. U.S. Const. art. 1, § 8, cl. 3.

52. Commerce ⚖︎13.5, 56

For laws that do not purposefully discriminate against out-of-state competitors but nevertheless incidentally burden interstate commerce, courts apply a more permissive balancing test for assessing whether the dormant Commerce Clause is violated, under which a state law will be struck down only if the burden imposed on interstate commerce clearly exceeds the putative local gains; the analysis primarily serves as an important reminder that a law's practical effects may also disclose the presence of a discriminatory purpose, and the test is most frequently deployed to detect the presence or absence of latent economic protectionism. U.S. Const. art. 1, § 8, cl. 3.

53. Commerce ⚖︎13.5

A state law that appears genuinely nondiscriminatory may still violate the dormant Commerce Clause if its burdens on commerce clearly outweigh the benefits of a state or local practice. U.S. Const. art. 1, § 8, cl. 3.

54. Commerce ⚖︎62.12, 81.5
    Trademarks ⚖︎1714(4)

Technology company failed to demonstrate that the New York statute providing injunctive relief from an injury from dilution of a distinctive trademark discriminated against out-of-state commerce, as would support finding that news organizations' claims for trademark dilution under the statute violated the dormant Commerce Clause; statute only provided injunctive relief for the dilution of a mark used in commerce in New York, and there was no indication that the statute facially discriminated against interstate commerce or harbored a discriminatory purpose, as there were no out-of-state competitors that would be disadvantaged vis-a-vis New York competitors due to the statute, and the statute did not articulate a discriminatory purpose against out-of-state competitors. U.S. Const. art. 1, § 8, cl. 3; N.Y. General Business Law §§ 360(a), 360(c), 360(h), 360-l.

55. Commerce ⚖︎62.12, 81.5
    Trademarks ⚖︎1714(4)

Technology company failed to demonstrate that the New York statute providing injunctive relief from an injury from dilution of a distinctive trademark substantially burdened interstate commerce in clear excess of its local benefits, as would support finding that news organizations' claims for trademark dilution under the statute violated the dormant Commerce Clause; there was no indication of latent economic protectionism in the statute. U.S. Const. art. 1, § 8, cl. 3; N.Y. General Business Law § 360-l.

56. Copyrights and Intellectual Property
    ⚖︎548, 554

To establish copyright infringement, the copyright owner must demonstrate that (1) the defendant has actually copied the plaintiff's work, and (2) the copying is illegal because a substantial similarity exists between the defendant's work and the protectible elements of plaintiff's. 17 U.S.C.A. § 501.

57. Copyrights and Intellectual Property
    ⚖︎559

The standard test in determining substantial similarity in a copyright infringement action is the ordinary observer test, which assesses whether an average lay observer would overlook any dissimilarities between the works and would conclude that one was copied from the other. 17 U.S.C.A. § 501.

58. Copyrights and Intellectual Property
    ⚖︎559

In determining whether substantial similarity exists between the work of the

defendant in a copyright infringement action and the protectible elements of plaintiff's work when the work contains both protectible and unprotectable elements, the ordinary observer test asks whether there exists substantial similarity between those elements, and only those elements, that provide copyrightability to the allegedly infringed work.   17 U.S.C.A. § 501.

59. Copyrights and Intellectual Property
     ☞1068(4)

The application of the ordinary observer test for determining whether substantial similarity exists between the work of the defendant in a copyright infringement action is by no means exclusively reserved for resolution by a jury; indeed, it is entirely appropriate for a district court to resolve that question as a matter of law, either because the similarity between two works concerns only non-copyrightable elements of the plaintiff's work, or because no reasonable jury, properly instructed, could find that the two works are substantially similar.   17 U.S.C.A. § 501.

60. Copyrights and Intellectual Property
     ☞292, 636

While facts themselves are not copyrightable, factual compilations may possess requisite originality for copyright protection; in such cases, where the compilation author clothes facts with an original collocation of words, he or she may be able to claim a copyright in this written expression, and others may copy the underlying facts from the publication, but not the precise words used to present them.   17 U.S.C.A. § 102(b).

61. Copyrights and Intellectual Property
     ☞636

When compiled facts are reported in a different arrangement, with a different sentence structure and different phrasing, the secondary work does not purloin protected expression, and no copyright infringement has ensued.   17 U.S.C.A. § 102(b).

62. Copyrights and Intellectual Property
     ☞636

Abridgements to first news organization's news articles created by technology company and artificial intelligence (AI) company, which partnered together to train AI large language models (LLM) using organization's news content, were not substantially similar to organization's articles, as required for organization's claim of direct copyright infringement; abridgments were detailed summaries, usually in bullet point form, of the facts contained in organization's articles, which such summaries differed in style, tone, length, and sentence structure from the articles, and the abridgements presented the facts in a different arrangement, such as bullet point lists or short summary paragraphs, with a different sentence structure and different phrasing.   17 U.S.C.A. § 501.

───────

Adnan Muttalib, Davida Brook, Ellie Rae Dupler, Emily K. Cronin, Susman Godfrey L.L.P., Los Angeles, CA, Alexander Frawley, Tamar Lusztig, Eudokia Spanos, Elisha Brandis Barron, Susman Godfrey L.L.P., New York, NY, Karen A. Chesley, The New York Times, New York, NY, Genevieve Vose Wallace, Ian B. Crosby, Katherine Marie Peaslee, Susman Godfrey L.L.P., Seattle, WA, Jennifer Maisel, Kristen Logan, Steven M. Lieberman, Rothwell Figg Ernst & Manbeck, P.C., Washington, DC, Scarlett Collings, Susman Godfrey LLP, Houston, TX, Zachary B. Savage, Susman Godfrey LLP, New York City, NY, for Plaintiff in 23-cv-11195.

Bryan Berkeley Thompson, Jennifer Maisel, Kristen Logan, Steven M. Lieber-

man, Jenny L. Colgate, Mark Rawls, Mary Mullins, Michael Jones, Robert Parker, Rothwell, Figg, Ernst & Manbeck, P.C., Washington, DC, Jeffrey A. Lindenbaum, Rothwell, Figg, Ernst & Manbeck, P.C., Ossining, NY, for Plaintiff in 24-cv-3285.

Jon Loevy, Kyle D. Wallenberg, Lauren Carbajal, Matthew Topic, Michael Kanovitz, Shelley Geiszler, Thomas Kayes, Stephen Stich Match, Steven Art, Loevy & Loevy Attorneys at Law, Chicago, IL, Steven M. Lieberman, Rothwell Figg Ernst & Manbeck PC, Washington, DC, for Plaintiff in 24-cv-4872.

Jeffrey S. Jacobson, Faegre Drinker Biddle & Reath LLP, New York, NY, Christopher Cariello, Orrick, Herrington & Sutcliffe LLP, New York, NY, Annette Louise Hurst, Orrick, Herrington & Sutcliffe LLP, San Francisco, CA, Paven Malhotra, Keker, Van Nest & Peters LLP, San Francisco, CA, Brianna Silverstein, Faegre Drinker Biddle & Reath LLP, Washington, DC, Carrie A. Beyer, Drinker Biddle & Reath, LLP, Chicago, IL, Elizabeth Mead Cavert Scheibel, Faegre Drinker Biddle & Reath LLP, Minneapolis, MN, Jared Barrett Briant, Kirstin Stoll-DeBell, Faegre Drinker Biddle & Reath LLP, Denver, CO, Laura Brooks Najemy, Sheryl Koval Garko, Orrick, Herrington & Sutcliffe LLP, Boston, MA, for Defendant Microsoft Corporation in 23-cv-11195.

Allison S. Blanco, Latham & Watkins LLP, Costa Mesa, CA, Allison Levine Stillman, Herman Heng Yue, Luke Budiardjo, Michael A. David, Rachel Renee Blitzer, Latham & Watkins LLP, New York, NY, Emily Claire Wood, Eric Nikolaides, Jocelyn Edith Greer, Morrison & Foerster LLP, New York, NY, Andrew Gass, Joseph Richard Wetzel, Jr., Latham & Watkins, LLP, San Francisco, CA, Daniel Jordan, Evangeline Zimmerman Burbidge, Lewis & Llewellyn LLP, San Francisco, CA, Joseph C. Gratz, Andrew L. Perito,

Morrison & Foerster LLP, San Francisco, CA, Robert A. Van Nest, R. James Slaughter, Andrew S. Bruns, Christopher S. Sun, Edward A. Bayley, Katie Lynn Joyce, Maile Nell Yeats-Rowe, Michelle S. Ybarra, Nicholas S. Goldberg, Paven Malhotra, Ryan Wong, Sarah Salomon, Thomas Edward Gorman, Andrew Dawson, Spencer McManus, Olivia Malone, Keker, Van Nest & Peters LLP, San Francisco, CA, Michael S. Kwun, Kwun Bhansali Lazarus LLP, San Francisco, CA, Elana Nightingale Dawson, Sarang Damle, Latham & Watkins LLP, Washington, DC, Carolyn M. Homer, Morrison Foerster, Washington, DC, Allyson R. Bennett, Rose Lee, Morrison & Foerster LLP, Los Angeles, CA, for Defendants OpenAI, Inc., OpenAI LP, OpenAI GP, LLC, OpenAI LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC, OpenAI Holdings, LLC in 23-cv-11195, 24-cv-3285.

Allison Levine Stillman, Herman Heng Yue, Michael A. David, Rachel Renee Blitzer, Latham & Watkins LLP, New York, NY, Emily Claire Wood, Eric Nikolaides, Morrison & Foerster LLP, New York, NY, Andrew Gass, Joseph Richard Wetzel, Jr., Latham & Watkins, LLP, San Francisco, CA, Daniel Jordan, Evangeline Zimmerman Burbidge, Lewis & Llewellyn LLP, San Francisco, CA, Joseph C. Gratz, Morrison & Foerster LLP, San Francisco, CA, Robert A. Van Nest, R. James Slaughter, Andrew S. Bruns, Christopher S. Sun, Edward A. Bayley, Katie Lynn Joyce, Michelle S. Ybarra, Nicholas S. Goldberg, Paven Malhotra, Ryan Wong, Thomas Edward Gorman, Andrew Dawson, Olivia Malone, Keker, Van Nest & Peters LLP, San Francisco, CA, Michael S. Kwun, Kwun Bhansali Lazarus LLP, San Francisco, CA, Elana Nightingale Dawson, Sarang Damle, Latham & Watkins LLP, Washington, DC, Carolyn M. Homer, Morrison Foerster, Washington,

DC, Allyson R. Bennett, Rose Lee, Morrison & Foerster LLP, Los Angeles, CA, for Defendants OpenAI, Inc., OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC, OpenAI Holdings, LLC in 24-cv-4872.

OPINION

SIDNEY H. STEIN, United States District Judge.

TABLE OF CONTENTS

Following plaintiffs' filing of complaints in the above-captioned copyright actions, defendants Microsoft Corporation and OpenAI Inc. et al. moved to dismiss several counts in all three actions. Specifically, defendants moved to dismiss (1) the contributory copyright infringement claims; (2) the Digital Millennium Copyright Act ("DMCA") claims; and (3) the common law unfair competition by misappropriation claims. Microsoft also moved to dismiss the state law trademark dilution claim in *Daily News LP, et al. v. Microsoft Corporation et al.*, No. 24-cv-3285 (the "*Daily News* action"), and OpenAI moved to dismiss (1) plaintiffs' direct copyright infringement claims involving conduct in 2019 and 2020 as time-barred by the statute of limitations; (2) plaintiffs' federal trademark dilution claim in the *Daily*

News action; and (3) the "abridgment" claims in *The Center for Investigative Reporting, Inc. v. OpenAI, Inc. et al.*, No. 24-cv-4872 (the "*CIR* action").

For the reasons that follow, the Court denies (1) OpenAI's motions to dismiss the direct infringement claims involving conduct occurring more than three years before the complaints were filed; (2) defendants' motions to dismiss the contributory copyright infringement claims; and (3) defendants' motions to dismiss the state and federal trademark dilution claims in the *Daily News* action.

The Court grants defendants' motions to dismiss the common law unfair competition by misappropriation claims and OpenAI's motion to dismiss the "abridgment" claims in the *CIR* action, and dismisses each of those claims with prejudice.

With respect to the DMCA claims, the Court grants (1) Microsoft's motions to dismiss the 17 U.S.C. § 1202(b)(1) claims against it in all three actions, (2) OpenAI's motion to dismiss the section 1202(b)(1) claim against it in *The New York Times Company v. Microsoft Corporation, et al.*, No. 23-cv-11195 (the "*Times* action"), and (3) defendants' motions to dismiss the section 1202(b)(3) claims against them in all three actions, and dismisses each claim without prejudice. The Court denies OpenAI's motions to dismiss the section 1202(b)(1) claims against it in the *Daily News* and *CIR* actions.

I. BACKGROUND

The following facts are taken from the plaintiffs' complaints and assumed to be true for the purpose of evaluating defendants' motions to dismiss under Federal Rule of Civil Procedure 12(b)(6). *See Faber v. Metro. Life Ins. Co.*, 648 F.3d 98, 104 (2d Cir. 2011).

A. Plaintiffs

Plaintiff The New York Times Company ("The Times") is a global, diversified multi-media company that publishes independent journalism through digital and print products, including its core news product The New York Times and other interest-specific publications including The Athletic, Cooking, Games, and Wirecutter. (*Times*, First Amended Complaint ("FAC") ¶ 14, ECF No. 170.) According to the *Times* complaint, The Times's journalism has garnered global recognition and won scores of accolades over the course of its more than 170-year existence for its high-quality, groundbreaking, and original reporting. (*Id.* ¶¶ 27–29.) This reporting—including investigative and breaking news, beat reporting, commentary and opinion, and in-depth reviews and analysis of arts and culture—covers numerous industries, topics and regions, and results from "an enormous amount of time, money, expertise, and talent." (*Id.* ¶¶ 26–37.) The Times has invested billions of dollars into its journalism to ensure it is accurate, independent, and fair. To support its resource-intensive reporting, The Times relies on subscription, advertising, licensing, and affiliate revenue. Since The Times launched its digital subscription plan and implemented its paywall in 2011—which requires payment for some, but not all, access to The Times's content—The Times has grown its paid digital and print subscribership to nearly 10.1 million subscribers, including approximately 50 to 100 million users engaging with its digital content each week. (*Id.* ¶¶ 41–45.) The Times owns more than 10 million registered, copyrighted works, which contain copyright management information ("CMI"), including title and other identifying information, copyright notice, terms and conditions of use, and identifying numbers or symbols referencing the CMI. (*Id.* ¶¶ 14, 125, 182.)

Plaintiffs in the *Daily News* action include Daily News, LP (the "New York Daily News"); Chicago Tribune Company, LLC, (the "Chicago Tribune"); Orlando Sentinel Communications Company, LLC (the "Orlando Sentinel"); Sun-Sentinel Company, LLC (the "Sun-Sentinel"); San Jose Mercury-News, LLC (the "Mercury News"); DP Media Network, LLC (the "Denver Post"); ORB Publishing, LLC (the "Orange County Register"); and Northwest Publications, LLC (the "Pioneer Press") (together, the "*Daily News* plaintiffs"). The *Daily News* plaintiffs collectively publish eight local newspapers across the United States, many of which have been in operation for more than 100 years and all of which have won Pulitzer Prizes and other national and local awards, providing critical local news coverage of many of the country's largest metropolitan areas to inform both local communities and the broader public. (*Daily News*, Compl. ¶¶ 40–47, ECF No. 1.) The publications are available in print and online and generate revenue from subscriptions, licensing, and advertising to help fuel the *Daily News* plaintiffs' billions of dollars of investments into the investigating and reporting of local news stories. (*Id.* ¶¶ 7, 48.) To protect and sustain their investment in local journalism, the *Daily News* plaintiffs keep some of their content behind a paywall, register their copyrights, and include copyright notices and other CMI in their publications. (*Id.* ¶ 49.)

Last, the Center for Investigative Reporting ("CIR") alleges in its complaint that it is the "oldest nonprofit newsroom in the country," whose "sole purpose is to benefit the public by reporting investigative stories about underrepresented voices in our democracy." (*CIR*, FAC ¶ 2, ECF No. 88.) CIR operates two relevant brands: Mother Jones, "a reader-supported news magazine and website known for ground-breaking investigative and in-

depth journalism on issues of national and global significance," and Reveal, which "operates an online news site" and "produces investigative journalism for the Reveal national public radio show and the Reveal podcast" that garners nearly 3 million podcast listeners per month. (*Id.* ¶¶ 13–14.) According to the *CIR* complaint, CIR spends significant time and money investigating complex stories and highlighting diverse issues and communities, and its brands have received awards for their "reporting, illustration, photography, videos, and social media." (*Id.* ¶ 13) CIR supports its reporting through licenses, advertising and affiliate revenue, as well as partnership agreements and programming. (*Id.* ¶ 3.) CIR also owns exclusive, registered copyrights to its Mother Jones magazine issues, as well as the works contained therein. (*Id.* ¶¶ 36–37.)

B. Defendants

Defendants are Microsoft Corporation ("Microsoft") and a web of interrelated entities including OpenAI Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global, LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC (collectively "OpenAI," and together with Microsoft, "defendants").

Founded in 2015 as a "non-profit artificial intelligence research company," OpenAI is now a commercial enterprise valued at roughly $90 billion as of the time the complaints were filed. (*Times*, FAC ¶¶ 55, 57.) OpenAI develops "large language models" or "LLMs." An LLM can receive text prompts as inputs and generate natural language responses as outputs, which result from the LLM's prediction of the most likely string of text to follow the inputted string of text based on its training on billions of written works. (*CIR*, FAC ¶¶ 48–49; *Times*, FAC ¶ 75; *Daily News*, Compl. ¶¶ 73–75.) OpenAI develops LLMs called Generative Pre-trained Transformers ("GPTs") and released the first version of its flagship GPT product, GPT-1, in 2018. (*Times*, FAC ¶ 58.) The release of GPT-2 followed in 2019, and both GPT-1 and GPT-2 were released on an open-source basis. (*Daily News*, Compl. ¶ 55; *Times*, FAC ¶ 58.) Starting in 2020 with the release of GPT-3, however, OpenAI "changed course": its GPT-3 model, along with its GPT-3.5 model (introduced in 2022) and GPT-4 model (introduced in 2023)—both of which were significantly more powerful than previous generations—were not released on an open-source basis. (*Daily News*, Compl. ¶ 56; *Times*, FAC ¶¶ 59, 83.) In November 2022, OpenAI released ChatGPT, a GPT-based, text-generating chatbot that, "given user-generated prompts, can mimic human-like natural language responses." (*Times*, FAC ¶ 61.) ChatGPT gained more than 100 million users within the first three months of its release. (*Id.*) OpenAI offers a free version of ChatGPT that is powered by GPT-3.5, as well as a premium service powered by GPT-4 for consumers who pay a $20 monthly subscription. (*Id.* ¶¶ 61–62; *Daily News*, Compl. ¶¶ 58–59.)

Defendant Microsoft has invested at least $13 billion in OpenAI Global LLC in exchange for receiving 75 percent of OpenAI Global's profits until Microsoft's investment is repaid, after which Microsoft will possess a 49 percent ownership stake in that company. (*Times*, FAC ¶ 15.) According to the complaints, Microsoft has "partnered with OpenAI deeply . . . for multiple years" in the "training, development, and commercialization of OpenAI's GPT products," including by providing and operating the cloud computing system OpenAI uses to train its models. (*Daily News*, Compl. ¶ 63; *Times*, FAC ¶ 66; *CIR*, FAC ¶ 26.) That cloud computing system was "specifically designed" for the purpose of "using essentially the whole internet,"

"in collaboration with and exclusively for OpenAI," "specifically to train that company's AI models." (*Daily News*, Compl. ¶¶ 66–67.)

Microsoft also collaborates with OpenAI to operate Microsoft's Copilot (formerly Bing Chat), a tool "designed to assist with the creation of documents, emails, presentations, and more." (*Times*, FAC ¶ 153; *Daily News*, Compl. ¶ 183; *CIR*, FAC ¶ 52.) Powered by GPT-4 and using Bing—Microsoft's internet search engine—Copilot responds to user queries in natural language to summarize content found on the internet. (*CIR*, FAC ¶ 52.) Finally, Microsoft and OpenAI have collaborated on "Browse with Bing," a plugin to ChatGPT released in May 2023 that also enables ChatGPT to access the latest content on the internet through the Microsoft Bing search engine. (*Times*, FAC ¶ 72.)

C. Defendants' Products

According to the complaints, defendants' LLMs implicate plaintiffs' works at two stages: (1) the training stage, where defendants use a corpus of text—including plaintiffs' works—to train their LLMs, and (2) the "output" stage, where defendants' LLMs generate outputs in response to user prompts that, according to the complaints, "regurgitate" plaintiffs' works. (*See Times*, FAC ¶¶ 77–79, 98.) Plaintiffs challenge at the output stage the outputs generated by (1) OpenAI's GPT products and (2) Microsoft's products powered by OpenAI's products. (*Id.* ¶ 118; *Daily News*, Compl. ¶ 114; *CIR*, FAC ¶ 81.) The Court will briefly discuss each process in turn.

1. The Training Stage

At the training stage, defendants first collect data, including plaintiffs' works, and then they train their LLMs on that data through a process that feeds the data through the model. (*Daily News*, Compl. ¶¶ 73, 80–81; *Times*, FAC ¶¶ 76, 83–84;

CIR, FAC ¶¶ 48–50; Jan. 14, 2025 Tr. of Oral Arg. at 7, ECF No. 433.) In particular, the collection stage (also known as the pre-training stage) involves collecting and storing a vast amount of content scraped from the internet, including content scraped from plaintiffs' websites, and creating datasets from that content which are later used to train the LLMs. (*Times* FAC ¶¶ 84–85; *Daily News*, Compl. ¶¶ 81–82.) Examples of these training datasets include (1) WebText and WebText2—developed using millions of links posted by "users of the 'Reddit' social network"—which OpenAI built and used to train GPT-2 and GPT-3 according to plaintiffs; and (2) Common Crawl, which is a "copy of the Internet" created by a third party. (*Times*, FAC ¶¶ 84–87; *Daily News*, Compl. ¶¶ 82–85; *CIR*, FAC ¶¶ 53–54, 70.) Plaintiffs allege that these datasets, among others used to train defendants' GPT models, contain a "staggering" amount of scraped content from plaintiffs' works, and that defendants have used and continue to use these and other training datasets to train their GPT models. (*See Times*, FAC ¶¶ 85–86.)

Next, defendants (1) "stor[e] copies of the training articles in computer memory," (2) encode the information from the training dataset in a numerical format, (3) "provid[e] a portion of the article to the model," and then (4) "adjust[ ] the parameters of the model so that the model accurately predicts the next word in the article." (*Daily News*, Compl. ¶ 75; *see also Times*, FAC ¶ 76; *CIR*, FAC ¶¶ 48, 50.) Defendants can further "fine tune" the models by "performing additional rounds of training using specific types of works to better mimic their content or style." (*Daily News*, Compl. ¶ 76.)

2. The "Output" Stage

The data that defendants collect at the pre-training stage and defendants' LLMs train on at the training stage inform the

responses of the LLMs to user queries at the output stage. In particular, LLMs respond to user queries by "predicting words that are likely to follow a given string of text based on the potentially billions of examples used to train [them]." (*Times*, FAC ¶ 75.) The result, according to the complaints, is that these outputs may "regurgitate" or reproduce large portions of plaintiffs' works, verbatim or nearly verbatim, that they have "memorized" during training in response to specific prompts. (*Times*, FAC ¶ 80; *Daily News*, Compl. ¶¶ 96, 144; *CIR*, FAC ¶¶ 81, 83–84.) Unfortunately, these outputs can also produce "hallucinations," which are output responses to user prompts that are "at best, not quite accurate and, at worst, demonstrably (but not recognizably) false." (*Times*, FAC ¶ 137.) According to plaintiffs, these hallucinations include outputs that misattribute content to plaintiffs that they did not in fact publish. (*See id.* ¶¶ 136–42; *Daily News*, Compl. ¶¶ 170–76.)

In addition, defendants also combine OpenAI's GPT-based technology with Microsoft's Bing search engine to search the internet and respond to user queries in natural language with the benefit of having access to the latest content on the internet. According to plaintiffs, these outputs generate "extensive paraphrases and direct quotes" of plaintiffs' works, without referring users to plaintiffs' websites in the same manner as do regular internet search engines, thereby obviating the need for users to visit plaintiffs' websites. (*Times* FAC ¶ 72; *Daily News* Compl. ¶ 69; *CIR*, FAC ¶ 52.)

D. The Actions

On December 27, 2023, The Times filed its complaint against Microsoft and OpenAI, seeking monetary and injunctive relief for (1) direct copyright infringement in violation of 17 U.S.C. § 501; (2) vicarious

copyright infringement; (3) contributory copyright infringement; (4) violations of the Digital Millennium Copyright Act ("DMCA"), 17 U.S.C. § 1202; (5) common law unfair competition by misappropriation; and (6) trademark dilution in violation of 15 U.S.C. § 1125(c). (*Times*, Compl., ECF No. 1.) In August 2024, The Times filed its First Amended Complaint, which included additional Times works but did not assert any new legal theories or causes of action. (*Times*, FAC, ECF No. 170.)

On April 30, 2024, the *Daily News* plaintiffs filed their complaint against defendants, also seeking monetary and injunctive relief and asserting the same claims as The Times, namely: direct copyright infringement, vicarious copyright infringement, contributory copyright infringement, DMCA violations, common law unfair competition by misappropriation, and federal trademark dilution. In addition, the *Daily News* plaintiffs asserted a claim of dilution and injury to business reputation in violation of New York General Business Law § 360-*l*. (*Daily News*, Compl., ECF No. 1.)

Finally, CIR filed a complaint against defendants in June 2024 and filed an amended complaint on September 24, 2024. (*CIR*, FAC, ECF No. 88.) As in the other two actions, CIR brings claims of direct copyright infringement, contributory copyright infringement, and DMCA violations, but does not raise the other claims that were included in the *Times* or *Daily News* complaints.

II. Legal Standard

"To survive a motion to dismiss, a complaint must contain sufficient factual matter, accepted as true, to 'state a claim to relief that is plausible on its face.'" *Ashcroft v. Iqbal*, 556 U.S. 662, 678, 129 S.Ct. 1937, 173 L.Ed.2d 868 (2009) (quoting *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570, 127 S.Ct. 1955, 167 L.Ed.2d 929 (2007)).

When considering a Rule 12(b)(6) motion to dismiss, a court must "draw all reasonable inferences in Plaintiffs' favor, assume all well-pleaded factual allegations to be true, and determine whether they plausibly give rise to an entitlement to relief." *Faber*, 648 F.3d at 104 (internal quotations marks and citation omitted). In so doing, the Court "is limited to facts stated on the face of the complaint and in documents appended to the complaint or incorporated in the complaint by reference, as well as to matters of which judicial notice may be taken." *Automated Salvage Transp., Inc. v. Wheelabrator Env't Sys., Inc.*, 155 F.3d 59, 67 (2d Cir. 1998); *see also Chambers v. Time Warner, Inc.*, 282 F.3d 147, 153–54 (2d Cir. 2002).

[1] Although the Court must accept as true all factual allegations, "[t]hreadbare recitals of the elements of a cause of action, supported by mere conclusory statements, do not suffice." *Iqbal*, 556 U.S. at 678, 129 S.Ct. 1937. By the same token, "[t]he choice between two plausible inferences that may be drawn from factual allegations is not a choice to be made by the court on a Rule 12(b)(6) motion. '[F]act-specific question[s] cannot be resolved on the pleadings.' " *Anderson News, L.L.C. v. Am. Media, Inc.*, 680 F.3d 162, 185 (2d Cir. 2012) (quoting *Todd v. Exxon Corp.*, 275 F.3d 191, 203 (2d Cir. 2001)).

III. DIRECT INFRINGEMENT OCCURRING MORE THAN THREE YEARS BEFORE THE FILING OF THE COMPLAINTS

OpenAI contends in its motions to dismiss the *Times* and *Daily News* complaints that plaintiffs' copyright claims involving conduct occurring more than three years prior to the filing of the complaints are time barred under 17 U.S.C. § 507(b).[1]

In particular, OpenAI asserts that plaintiffs' direct infringement claims based on OpenAI's creation and use of the GPT-2 and GPT-3 training datasets are time barred because the alleged infringements occurred more than three years before the filing of their complaints in December 2023 by The Times and April 2024 by the *Daily News* plaintiffs. Plaintiffs disagree and contend that OpenAI has not met its burden of establishing that The Times discovered the alleged infringement before December 27, 2020, three years before it filed its complaint, or that the *Daily News* plaintiffs discovered the alleged infringement before April 30, 2021, three years before they filed their complaint in the *Daily News* action. The Court agrees with plaintiffs.

A. Applicable Standard

[2] Section 507(b) states that "[n]o civil action shall be maintained under the provisions of this title unless it is commenced within three years after the claim accrued." 17 U.S.C. § 507(b). Under the discovery rule, "an infringement claim does not 'accrue' until the copyright holder discovers, or with due diligence should have discovered, the infringement." *Sohm v. Scholastic Inc.*, 959 F.3d 39, 50 (2d Cir. 2020), *abrogated on other grounds, Warner Chappell Music, Inc. v. Nealy*, 601 U.S. 366, 373, 144 S.Ct. 1135, 218 L.Ed.2d 363 (2024).

[3–5] The discovery rule does not impose on copyright holders "a general duty to police the internet" to uncover infringement. *Parisienne v. Scripps Media, Inc.*, No. 19-cv-8612, 2021 WL 3668084, at *2 (S.D.N.Y. Aug. 17, 2021) (internal quotation marks omitted). Indeed, the Court of

---

1. In this Section III, the Court uses the term "plaintiffs" to refer to The Times and the *Daily News* plaintiffs exclusively.

Appeals for the Second Circuit has rejected the argument that a plaintiff's failure to conduct a search to uncover potential infringement, despite having the ability to do so, alone triggers constructive notice. *See Sohm*, 959 F.3d at 51. Similarly, "a copyright holder's *general* diligence or allegations of diligence in seeking out and litigating infringements, alone, are insufficient to make it clear that the holder's *particular* claims in any given case should have been discovered more than three years before the action's commencement." *Michael Grecco Prods., Inc. v. RADesign, Inc.*, 112 F.4th 144, 148 (2d Cir. 2024). Rather, to establish constructive notice, a defendant must identify specific "facts or circumstances that would have prompted such an inquiry" by the copyright holder into the alleged infringing activity. *Sohm*, 959 F.3d at 51; *see also McGlynn v. Sinovision Inc.*, No. 23-cv-4826, 2024 WL 643021, at *2 (S.D.N.Y. Feb. 15, 2024).

[6, 7] Finally, "[t]here is no 'sophisticated plaintiff' exception to the discovery rule." *RADesign*, 112 F.4th at 148. "The date on which a copyright holder, with the exercise of due diligence, would have discovered an infringement—or whether the alleged date of discovery reflected a lack of due diligence—is a fact-intensive inquiry that cannot be determined from the general nature of a copyright holder's 'sophistication' alone." *Id.* at 152.

[8] Because the statute of limitations is an affirmative defense, OpenAI bears the burden of establishing that by December 27, 2020 for The Times and April 30, 2021 for the *Daily News* plaintiffs—three years before they filed their respective complaints—the plaintiffs should have been aware of the alleged infringement. *See id.* at 149; Fed. R. Civ. P. 8(c)(1). Dismissal of a copyright infringement claim on statute

of limitations grounds at the pleadings stage is only appropriate when "it is clear from the face of the complaint, and matters of which the court may take judicial notice, that the plaintiff's claims are barred as a matter of law." *Sewell v. Bernardin*, 795 F.3d 337, 339 (2d Cir. 2015) (citation omitted). "However, where there is even 'some doubt' as to whether dismissal is warranted, a court should not grant a Rule 12(b)(6) motion on statute of limitations grounds." *PK Music Performance, Inc. v. Timberlake*, No. 16-cv-1215, 2018 WL 4759737, at *7 (S.D.N.Y. Sept. 30, 2018) (citing *Ortiz v. Cornetta*, 867 F.2d 146, 149 (2d Cir. 1989)).

B. Plaintiffs' Claims of Direct Infringement Occurring More than Three Years Before the Filing of the Complaints Are Not Time-Barred by the Statute of Limitations.

[9] OpenAI has not met its burden of establishing that The Times and the *Daily News* plaintiffs discovered, or with due diligence should have discovered, the alleged infringement before December 27, 2020 and April 30, 2021 respectively. Although the complaints allege that defendants trained their LLMs in 2019 and 2020 on datasets that included plaintiffs' works, the complaints do not establish that plaintiffs "discover[ed], or with due diligence should have discovered" that fact in 2019 and 2020. *See Sohm*, 959 F.3d at 50.

OpenAI identifies a few publicly available documents from 2019 and 2020, one of which—an article by Jennifer Langston entitled "Microsoft Announces New Supercomputer, Lays Out Vision for Future AI Work"—was cited in plaintiffs' complaints, to argue that it was "common knowledge" by 2020 that some of defendants' training

datasets included plaintiffs' works.[2] These documents are insufficient for at least two reasons. First, as to the Langston article, OpenAI does not explain why plaintiffs would have known in 2020 of that article's existence, nor does it point to "facts or circumstances" that would have prompted plaintiffs to look for the article at that time. *See Sohm*, 959 F.3d at 51. Second, OpenAI fails to explain why the articles, even if their existence had been known to plaintiffs at the time of their publishing, are sufficient to put plaintiffs on notice of the particular infringing conduct by defendants that provides the basis for plaintiffs' claims. *Cf. McGlynn*, 2024 WL 643021, at *5 (declining to infer that past litigation between the parties constituted "facts or circumstances" sufficient to put copyright holder on notice of alleged infringement).

OpenAI makes much of The Times's reporting in November 2020 that OpenAI trained its models by "analyzing . . . nearly a trillion words posted to blogs, social media and the rest of the internet." Cade Metz, *Meet GPT-3. It Has Learned To Code (and Blog and Argue)*, N.Y. Times (Nov. 24, 2020), https://www.nytimes.com/2020/11/24/science/artificial-intelligence-ai-gpt3.html. The fact that one of The Times's reporters discussed OpenAI's "analyzing . . . a trillion words" on the internet fails to "make it clear that [plaintiffs']

particular claims . . . should have been discovered more than three years before the action's commencement." *RADesign*, 112 F.4th at 148. Those claims involve the specific copying of plaintiffs' works by OpenAI, which, as alleged in the complaints, "became a household name upon the release of ChatGPT in November 2022," two years after the Metz article. (*Times*, FAC ¶ 61; *Daily News*, FAC ¶ 58.)

Finally, OpenAI's argument that The Times, as a "sophisticated publisher," had a duty "to take prompt action after being put on notice of what it now claims to be alleged infringement" is a straw man. (*See Times*, ECF No. 75 at 3.) OpenAI has failed to establish that The Times *was in fact on notice* before December 27, 2020—and the Second Circuit has squarely rejected a heightened "sophisticated rights-holder" theory of constructive knowledge. *RADesign*, 112 F.4th at 148.

Discovery may reveal facts supporting OpenAI's contention that The Times and the *Daily News* plaintiffs discovered the alleged infringement before December 27, 2020 and April 30, 2021 respectively. At the motion to dismiss stage, however, OpenAI's conclusory statement that plaintiffs "discovered or with reasonable diligence should have discovered these activities"

---

2. *See, e.g., Daily News*, Compl. ¶ 67 n.11 (citing Jennifer Langston, *Microsoft Announces New Supercomputer, Lays Out Vision for Future AI Work*, Microsoft (May 19, 2020), https://news.microsoft.com/source/features/ai/openai-azure-supercomputer/) (article published on Microsoft's website discussing a "new class of multitasking AI models" that "can learn about language by examining billions of pages of publicly available documents on the internet"); *Times*, FAC ¶ 70 n.8 (same). *See also Times*, ECF No. 52 at 6 (citing Cade Metz, *Meet GPT-3. It Has Learned To Code (and Blog and Argue)*, N.Y. Times (Nov. 24, 2020), https://www.nytimes.com/2020/11/science/artificial-intelligence-ai-gpt3.html).

OpenAI references the Metz article even though it was not cited in The Times's complaint, contending that the article is not being introduced for the truth of its statements but for the Court to "take judicial notice of the fact [of] press coverage . . . [to] decid[e] whether so-called 'storm warnings' were adequate to trigger inquiry notice." (*Times*, ECF No. 75 at 2 n.2) (quoting *Staehr v. Hartford Fin. Servs. Grp., Inc.*, 547 F.3d 406, 425 (2d Cir. 2008).) The Court takes judicial notice of the Metz article but finds that its publication does not constitute a "storm warning . . . adequate to trigger inquiry notice." *Staehr*, 547 F.3d at 425.

prior to three years before the filing of their complaints (*Times*, ECF No. 52 at 15 n.33), fails to meet its burden of establishing actual or constructive knowledge. Accordingly, OpenAI's motion to dismiss the copyright infringement claims arising more than three years before The Times and the *Daily News* plaintiffs filed their complaints is denied.

IV. CONTRIBUTORY COPYRIGHT INFRINGEMENT

Next, defendants move to dismiss plaintiffs' contributory copyright infringement claims. Plaintiffs bring those claims in the alternative to their direct infringement claims to the extent third-party end users—not defendants—are found liable for direct infringement for generating infringing outputs using defendants' LLMs. Under this theory, defendants materially contributed to and directly assisted with the direct infringement by end users by (1) building and training their LLMs using plaintiffs' works; (2) deciding what content is outputted by their LLMs through specific training techniques; and (3) developing LLMs capable of distributing copies of plaintiffs' works to end users without authorization by plaintiffs.

Defendants contend that plaintiffs have failed to plausibly allege both direct infringement by a third party and that defendants knew of third-party infringement.

A. Applicable Standard

[10, 11] An individual or entity may be held liable as a contributory copyright infringer if that individual or entity, "with knowledge of the infringing activity, induces, causes or materially contributes to the infringing conduct of another." *Gershwin Publ'g Corp. v. Columbia Artists Mgmt., Inc.*, 443 F.2d 1159, 1162 (2d Cir. 1971); *see also Arista Recs., LLC v. Doe 3*, 604 F.3d 110, 117 (2d Cir. 2010). To plau-

sibly allege contributory copyright infringement under a theory of material contribution, a plaintiff must show "(1) direct infringement by a third party, (2) that the defendant had 'knowledge of the infringing activity,' (3) and that the defendant 'materially contribute[d] to' the third party's infringement." *Dow Jones & Co., Inc. v. Juwai Ltd.*, No. 21-cv-7284, 2023 WL 2561588, at *3 (S.D.N.Y. Mar. 17, 2023) (citing *Smith v. BarnesandNoble.com, LLC*, 143 F. Supp. 3d 115, 124 (S.D.N.Y. 2015)).

The parties disagree on the standard for establishing that defendants knew a third party was infringing plaintiffs' copyrights. Their disagreement mirrors a split among the Circuits regarding the scienter required to support a contributory copyright infringement claim, absent actual knowledge of third-party infringement. Plaintiffs contend that the standard is actual or constructive knowledge; namely, whether defendants objectively "know or have reason to know" of the direct infringement by third-party end users. That is the standard in the Second Circuit. *See Gershwin Publ'g Corp.*, 443 F.2d at 1162; *Doe 3*, 604 F.3d at 118; *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640, 658 (S.D.N.Y. 2013); *State St. Glob. Advisors Tr. Co. v. Visbal*, 431 F. Supp. 3d 322, 358 (S.D.N.Y. 2020); *Rams v. Def Jam Recordings, Inc.*, 202 F. Supp. 3d 376, 383 (S.D.N.Y. 2016); *BarnesandNoble.com*, 143 F. Supp. 3d at 124; *Arista Recs. LLC v. Usenet.com, Inc.*, 633 F. Supp. 2d 124, 154 (S.D.N.Y. 2009); *Arista Recs., Inc. v. Mp3Board, Inc.*, No. 00-cv-4660, 2002 WL 1997918, at *6–7 (S.D.N.Y. Aug. 29, 2002).

Defendants urge a heightened standard, contending that liability for contributory copyright infringement requires that the defendant have possessed actual knowledge of or willful blindness to specific acts of infringement. That is the standard in

the Ninth Circuit. *See Ludvarts, LLC v. AT&T Mobility, LLC*, 710 F.3d 1068, 1072–73 (9th Cir. 2013) (requiring "actual knowledge of specific acts of infringement" or "[w]illful blindness of specific facts" to "establish knowledge for contributory [copyright] liability"). The Court therefore will evaluate plaintiffs' contributory copyright infringement claims under the standard in the Second Circuit.

[12–14] "The knowledge standard is an objective one; contributory infringement liability is imposed on persons who 'know *or have reason to know*' of the direct infringement." *Doe 3*, 604 F.3d at 118 (citation omitted). "[M]ore than a generalized knowledge by the defendant of the possibility of infringement" is required to meet the knowledge requirement. *Hartmann v. Apple, Inc.*, No. 20-cv-6049, 2021 WL 4267820, at *6 (S.D.N.Y. Sept. 20, 2021) (citing *Rams*, 202 F. Supp. 3d at 376). By the same token, "knowledge of specific infringements is not required to support a finding of contributory infringement." *Usenet.com, Inc.*, 633 F. Supp. 2d at 154. Rather, courts have looked to whether the complaint contains "allegation[s] that the defendant investigated or would have had reason to investigate the alleged infringement," *Hartmann v. Popcornflix.com LLC*, 690 F. Supp. 3d 309, 320 (S.D.N.Y. 2023), as well as evidence of "cease-and-desist letters, officer and employee statements, promotional materials, and industry experience" to determine whether the defendant was put on notice of third-party infringement. *ReDigi Inc.*, 934 F. Supp. 2d at 658.

[15] To establish that the defendant "materially contributed" to the infringement, the complaint must show that the defendant "encouraged or assisted others' infringement[ ] or provided machinery or goods that facilitated infringement." *Arista Recs. LLC v. Lime Grp. LLC*, 784 F. Supp. 2d 398, 432 (S.D.N.Y. 2011).[3]

B. Plaintiffs Have Plausibly Alleged Contributory Copyright Infringement.

1. Third Party Infringement

[16] Although defendants rely primarily on cases decided on summary judgment or after trial, the question at this stage is simply whether plaintiffs have plausibly alleged that end-user infringement has taken place. The Court finds that plaintiffs have done so. First, the complaints allege "widely publicized" instances "of copyright infringement after ChatGPT, Browse with Bing, and Bing Chat were released." (*Times*, FAC ¶ 126; *Daily News*, Compl. ¶ 141.) Second, plaintiffs include numerous examples of infringing outputs in their complaints. (*See, e.g., Times*, FAC ¶¶ 99–101, 104–22, Ex. J; *Daily News*, Compl. ¶¶ 98–113, 118–37, Ex. J; *CIR*, FAC Ex. 10, Ex. 11 at 6–17, 14.)[4] These examples "raise a reasonable expectation" that discovery will reveal evidence of additional examples of third-party infringement such that dismissal at this stage would be improper. *See Twombly*, 550 U.S. at 556, 127 S.Ct. 1955.

Matthew Bender v. West Publishing Co., relied upon by defendants, does not counsel a different result. *See* 158 F.3d 693 (2d Cir. 1998). In that case, which was decided

---

3. Defendants do not challenge the "material contribution" component of plaintiffs' contributory copyright infringement claim.

4. The *Daily News* and *Times* complaints include numerous examples of allegedly infringing outputs. The *CIR* complaint includes five

examples of allegedly infringing outputs, of which at least two illustrate the ease with which end-user infringement can occur using defendants' products. (*See CIR*, FAC Ex. 10, Ex. 11 at 6–7, 14.)

on a motion for summary judgment, the rightsholder (West) held a copyright in its printed compilations of judicial opinions, which were distinctly paginated in the star pagination format. *Id.* at 696. A CD-ROM disc manufacturer (Matthew Bender) produced CD-ROM discs containing compilations of judicial opinions that included references to West's star pagination, but did not arrange the judicial opinions as West had done in its printed compilations. *Id.* at 697. West alleged that Matthew Bender was liable for the copyright infringement of third-party users. In affirming the district court's grant of summary judgment for Matthew Bender, the Second Circuit determined that West had failed to plausibly allege the existence of third-party infringement, reasoning that to infringe West's copyright in its printed compilation of court opinions, "a user must retrieve each case, one at a time, in the order in which they appear in the West volume, and then print each one." *Id.* at 706. The court concluded that West had failed to identify any third-party infringer other than its own counsel (who had artificially replicated West's compilation of cases) and that West's "hypothesized" examples of infringement were insufficient. *Id.*

In contrast, plaintiffs' examples of allegedly infringing outputs at the pleading stage—including more than 100 pages of examples provided in Exhibit J to the *Times* complaint, and dozens of examples in Exhibit J to the *Daily News* complaint—combined with their allegations of "widely publicized" instances of copyright infringement by end users of defendants' products, give rise to a plausible inference of copyright infringement by third parties.

2. Knowledge of Third-Party Infringing Activity

[17] Plaintiffs have also plausibly alleged that defendants possessed constructive, if not actual, knowledge of end-user infringement. The complaints allege that defendants knew they were using copyrighted works to train their models and were fully aware of plaintiffs' protected interests in their works. (*See Times*, FAC ¶¶ 7, 8, 69–70, 126; *Daily News*, Compl. ¶¶ 66–67, 114–15, 144–47; *CIR*, FAC ¶¶ 26, 28, 83, 144.) *Cf. Hartmann v. Amazon.com, Inc.* No. 20-cv-4928, 2021 WL 3683510, at *7 (S.D.N.Y. Aug. 19, 2021) (reasoning that defendants did not know or have reason to know of third-party infringement in part because defendants did not know plaintiff had a protected interest in the underlying works). The complaints also reference "widely publicized" reports that end users were using defendants' LLMs to "elicit copyrighted content" (*Times*, FAC ¶ 126; *Daily News*, Compl. ¶ 141), and statements by OpenAI representatives about internal company disagreements regarding copyright issues. (*Times*, FAC ¶ 124; *Daily News*, Compl. ¶ 139.) The Times even informed defendants "that their tools infringed its copyrighted works," supporting the inference that defendants possessed actual knowledge of infringement by end users. (*Times*, FAC ¶ 126.) Taken as true, these facts give rise to a plausible inference that defendants at a minimum had reason to investigate and uncover end-user infringement.

Plaintiffs allege that defendants possessed far more than a "generalized knowledge of the possibility" of third-party infringement. *Visbal*, 431 F. Supp. 3d at 358. Indeed, plaintiffs allege both that (1) defendants possessed actual knowledge of third-party infringement and (2) defendants knew not only that their unauthorized copying of plaintiffs' works on a massive scale during the training of their LLMs would "result[ ] in the unauthorized encoding of huge numbers of such works in the models themselves," but also that it

"would inevitably result in the unauthorized display of such works" in response to third-party queries. (*Times*, FAC ¶ 124; *Daily News*, Compl. ¶ 139.) In other words, defendants knew or had reason to know of third-party infringement because copyright infringement was "central to [defendants'] business model." *See ReDigi*, 934 F. Supp. 2d at 659; *see also Andersen v. Stability AI Ltd.*, 744 F. Supp. 3d 956, 969 (N.D. Cal. 2024) (finding defendants possessed knowledge of third-party infringement when their products were "built to a significant extent on copyrighted works" and their operation "necessarily invokes copies or protected elements of those works"). These allegations are sufficient at the pleading stage to establish a plausible inference that defendants possessed actual or constructive knowledge of third-party infringement.

3. Substantial Noninfringing Uses

Defendants also contend that plaintiffs' contributory copyright infringement claims fail because defendants' LLMs—even if they were used by third parties to commit copyright infringement, and even if defendants had knowledge of that fact—are capable of "substantial noninfringing uses." *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 442, 104 S.Ct. 774, 78 L.Ed.2d 574 (1984); *see also Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 937, 125 S.Ct. 2764, 162 L.Ed.2d 781 (2005). Defendants rely on the U.S. Supreme Court's decisions in *Sony* and *Grokster* for this position, but that reliance is misplaced. In *Sony*—which involved Sony's video tape recorder (VTR)—the Supreme Court held that "[t]he sale of copying equipment, like the sale of other articles of commerce, does not constitute contributory infringement if the product is widely used for legitimate, unobjectionable purposes." *Sony*, 464 U.S. at 442, 104 S.Ct. 774. However, in *Grokster*—which involved

peer-to-peer file sharing computer software—the Supreme Court clarified that a defendant whose product is capable of substantial noninfringing can still be held liable for third-party infringement in certain circumstances. *Grokster*, 545 U.S. at 933–35, 125 S.Ct. 2764. In that case, the Supreme Court held that a distributor of a device who "promot[es] its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement," is liable for third-party infringement. *Id.* at 919, 125 S.Ct. 2764.

Sony and *Grokster* do not foreclose plaintiffs' contributory copyright infringement claims at this stage for at least three reasons. First, *Sony* and *Grokster* involved cases decided either on summary judgment or after trial rather than on motions to dismiss the complaint. Indeed, many of the decisions relied upon by defendants to support their contention that their LLMs have "substantial noninfringing uses" were decided either on summary judgment or after trial. *See, e.g., BarnesandNoble.com*, 143 F. Supp. 3d at 124; *Matthew Bender*, 158 F.3d at 706–07. The question before the Court on a motion to dismiss a contributory copyright infringement claim is narrow: whether plaintiffs have plausibly alleged that defendants knew or had reason to know of actual third-party infringement by end users of their products. *See Doe 3*, 604 F.3d at 117–18.

Second, while *Sony* and *Grokster* analyzed claims of contributory copyright infringement by inducement, they did not discuss claims of contributory copyright infringement by material contribution, which plaintiffs allege here. *See Grokster*, 545 U.S. at 934, 125 S.Ct. 2764 (rejecting circuit court's "converting [of] the [*Sony*] case from one about liability resting on imputed intent to one about liability on any theory"). Indeed, "the fact that a product is 'capable of substantial lawful use' does

not mean the 'producer can never be held contributorily liable.'" *BMG Rts. Mgmt. (US) LLC v. Cox Commc'ns, Inc.*, 881 F.3d 293, 306 (4th Cir. 2018) (citations omitted).

In a word, *Sony* foreclosed imputing "culpable intent" solely based on the "characteristics or uses of a distributed product." *Grokster*, 545 U.S. at 934, 125 S.Ct. 2764. It left open, however, other "rules of fault-based liability derived from the common law," *id.*, including liability based on material contribution, a theory which neither *Sony* nor *Grokster* discussed nor foreclosed. *See Gershwin Publ'g Corp.*, 443 F.2d at 1162–63 (predicating contributory liability for material contribution on "the common law doctrine that one who knowingly participates [in] or furthers a tortious act is jointly and severally liable with the prime tortfeasor").

Third, the facts in *Sony* did not include two important distinguishing features. First, in *Sony* there was no "ongoing relationship between the direct infringer and the contributory infringer at the time the infringing conduct occurred." *Sony*, 464 U.S. at 437, 104 S.Ct. 774. Here, however, an "ongoing relationship" exists between defendants and end users, via defendants' LLM outputs that respond to end users' prompts. *Cf. id.* at 438, 104 S.Ct. 774 ("The only contact between Sony and the users of [the VTRs] occurred at the moment of sale."). Second, the VTR was not a product that itself was *built* on purportedly appropriated works, as are defendants' products here. *Cf. id.* at 421, 104 S.Ct. 774 (describing defendant as a distributor of "copying equipment"). On the second point, *Sony*'s discussion of *Kalem Co. v. Harper Brothers*, 222 U.S. 55, 32 S.Ct. 20, 56 L.Ed. 92 (1911) is instructive. In *Kalem*, the Supreme Court "held that the produc-

er of an unauthorized film dramatization of the copyrighted book *Ben Hur* was liable for his sale of the motion picture to jobbers, who in turn arranged for the commercial exhibition of the film." *Sony*, 464 U.S. at 435, 104 S.Ct. 774 (discussing *Kalem*). Unlike the defendant in *Sony*, the defendant producer in *Kalem* "did not merely provide the 'means' to accomplish an infringing activity; the producer supplied the work itself, albeit in a new medium of expression." *Id.* at 436, 104 S.Ct. 774. This form of contributory infringement is analogous to the material contribution by defendants that plaintiffs allege. According to the complaints, defendants appropriated plaintiffs' works, created the "tangible medium" upon which those protected works were "recorded," *id.*, and provided third-party infringers with both the means of infringing and the works from which to do so.

* * *

The Court finds that plaintiffs have plausibly alleged the existence of third-party end-user infringement and that defendants knew or had reason to know of that infringement. Accordingly, plaintiffs have plausibly alleged their contributory copyright infringement claims and defendants' motions to dismiss those claims in all three actions are denied.

V. THE DMCA CLAIMS

Plaintiffs in each action bring two claims under the Digital Millennium Copyright Act ("DMCA") against Microsoft and OpenAI. The first claim is brought pursuant to 17 U.S.C. § 1202(b)(1), which prohibits "intentionally remov[ing] or alter[ing] any copyright management information" ("CMI").[5] The second claim is brought pur-

---

5. CMI comprises information "conveyed in connection with copies … of a work," including, as relevant here, "[t]he title and oth-

er information identifying the work," "the author of the work," and "[t]he name of, and

suant to 17 U.S.C. § 1202(b)(3), which prohibits the "distribution" of "works" or "copies of works … knowing that [CMI] has been removed or altered without authority of the copyright owner." In both provisions, the defendant must also have "know[n]" or "ha[d] reasonable grounds to know" that its conduct would "induce, enable, facilitate, or conceal an infringement." 17 U.S.C. § 1202(b). Section 1202(b) claims therefore contain a "double-scienter" requirement: the defendant must (1) "intentionally" remove CMI under section 1202(b)(1) or distribute copyrighted works "knowing" that CMI was removed under section 1202(b)(3); and (2) know or have reason to know that its conduct would induce, enable, facilitate, or conceal infringement. *See Mango v. BuzzFeed, Inc.*, 970 F.3d 167, 171 (2d Cir. 2020).

Defendants move to dismiss both DMCA claims, contending that plaintiffs lack Article III and statutory standing. They also contend that the claims fail on the merits, because plaintiffs have failed to plausibly allege (1) the actual removal of CMI; (2) that defendants had knowledge that the removal would lead to infringement; and (3) that defendants "distributed" "copies" of plaintiffs' works.

A. Article III & Statutory Standing
1. Article III Standing

[18, 19] In order to establish Article III standing, the U.S. Constitution requires a plaintiff to 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, 141 S.Ct. 2190, 210 L.Ed.2d 568 (2021); U.S. Const. art. III, § 2, cl. 1. To be

"concrete," the injury must be "actual or imminent, not conjectural or hypothetical." *Spokeo, Inc. v. Robins*, 578 U.S. 330, 339, 136 S.Ct. 1540, 194 L.Ed.2d 635 (2016) (quoting *Lujan v. Defs. of Wildlife*, 504 U.S. 555, 560, 112 S.Ct. 2130, 119 L.Ed.2d 351 (1992)).

[20, 21] While "Congress may 'elevate' harms that 'exist' in the real world before Congress recognized them to actionable legal status, it may not simply enact an injury into existence, using its lawmaking power to transform something that is not remotely harmful into something that is." *TransUnion*, 594 U.S. at 426, 141 S.Ct. 2190 (citation omitted). Regardless of the existence *vel non* of a statutory cause of action, courts must "independently decide whether a plaintiff has suffered a concrete harm under Article III." *Id.* To do so, courts look to "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," namely "whether plaintiffs have identified a close historical or common-law analogue for their asserted injury." *Id.* at 424, 141 S.Ct. 2190 (quoting *Spokeo*, 578 U.S. at 341, 136 S.Ct. 1540).

Plaintiffs allege that defendants removed CMI from plaintiffs' works during the process of training their LLMs and in distributing unauthorized copies of plaintiffs' works through regurgitating outputs. (*See Times*, FAC ¶¶ 184–86; *Daily News*, Compl. ¶ 159; *CIR*, FAC ¶ 103.) According to plaintiffs, this removal caused two forms of concrete harm: (1) harm caused through the removal of CMI during the training process, absent dissemination, and (2) harm caused by the dissemination of CMI-less copies of plaintiffs' works in the form of LLM outputs.

---

other identifying information about, the copy-

right owner of the work." 17 U.S.C. § 1202(c).

Defendants contend that both theories of harm fail. With respect to the first theory—harm absent dissemination—defendants contend that because the training data from which CMI was allegedly removed was never disseminated or otherwise made publicly available, it is therefore not a legally cognizable injury under Article III. The second theory of harm—harm caused by dissemination—fails for two reasons, according to defendants. First, the alleged injuries caused by the dissemination of CMI-less works—including the inability to receive licensing and subscription revenue, and the possibility that the regurgitating outputs will divert readers from plaintiffs' platforms—do not have any nexus to CMI removal. Second, because the outputs cited in the complaints either reference plaintiffs' articles by name or result from prompts that quote substantial portions of the underlying article, "any user who encountered those outputs would have no doubt as to the provenance of the text and could easily find it on the [plaintiffs'] website," thereby making plaintiffs' alleged injuries imaginary and not "concrete." (*Times*, ECF No. 52 at 22; *see also Daily News*, ECF No. 82 at 13–14; *CIR*, ECF No. 100 at 17.)

2. Statutory Standing

OpenAI also raises a statutory standing argument, contending that plaintiffs have failed to allege that they are "person[s] injured by" a DMCA violation, as required by 17 U.S.C. § 1203(a). Section 1203(a) states that "[a]ny person injured by a violation of section 1201 or 1202 may bring a civil action in an appropriate United States district court for such violation." OpenAI contends that the language of this section requires an injury *beyond* a mere statutory violation, and that the injury must be caused by the section 1202(b) violation specifically. It also contends that, for the same reasons plaintiffs have failed to allege an injury for Article III purposes, plaintiffs fail to establish that it was either the *CMI removal* from defendants' training datasets or the dissemination of *CMI-less* works that caused plaintiffs the specific harm they allege.

3. Plaintiffs Have Article III Standing To Bring Their DMCA Claims.

a. Concreteness

[22] Defendants challenge plaintiffs' allegations of harm as lacking "a close relationship to harms traditionally recognized as providing a basis for lawsuits in American courts," relying on the Supreme Court's decision in *TransUnion LLC v. Ramirez.* 594 U.S. at 425, 141 S.Ct. 2190. In *TransUnion*, the Supreme Court explained that the harm alleged must possess a "close historical or common-law analogue," although, importantly, that harm need not be an "exact duplicate." *Id.* at 424, 141 S.Ct. 2190. Here, traditional copyright law provides that "close historical or common-law analogue" and supports plaintiffs' claims of harm.

"Copyright claims predate the Constitution's ratification." *The Intercept Media, Inc. v. OpenAI, Inc.*, No. 24-cv-1515, 767 F.Supp.3d 18, 26 (S.D.N.Y. Feb. 20, 2025) (citing The Federalist No. 43 (James Madison)). Indeed, copyright is listed among Congress' enumerated powers in Article I of the Constitution, *see* U.S. Const. art. I, § 8, cl. 8, and, from the time of the enactment of the Copyright Act of 1790, Congress has updated the copyright laws numerous times over the past two centuries. *See* Copyright Act of 1790; Copyright Act of 1831; Copyright Act of 1870; Copyright Act of 1909; Copyright Act of 1976.

[23, 24] DMCA claims differ from traditional copyright claims. DMCA claims protect against harms caused by the unauthorized removal of CMI from a copyright-

ed work; traditional copyright infringement claims protect against, among other things, the unauthorized reproduction and distribution of protected works. *See Mango*, 970 F.3d at 170–71; *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 95 (2d Cir. 2014). Nonetheless, copyright infringement—a harm "traditionally recognized as providing a basis for lawsuits in American Courts"—provides an appropriate "close historical or common-law analogue" to the harm caused by a DMCA violation. *TransUnion*, 594 U.S. at 424–25, 141 S.Ct. 2190. Both traditional copyright and DMCA claims are grounded in notions of property rights, and both claims are designed to "promote the Progress of Science and useful Arts." U.S. Const. art. I, § 8, cl. 8. As Judge Jed S. Rakoff wrote in *The Intercept Media*, "[t]he DMCA adds another stick to the bundle of property rights already guaranteed to an author in her work under traditional copyright law," and "[t]he fact that the specific right at issue here is not expressly rooted in that overall history misses the point; the exact contours of the property rights given to a copyright holder are not frozen in time by the Copyright Act of 1790." 767 F.Supp.3d at 27.

As noted above, Article III requires that the *harm* alleged have a "close relationship to a harm traditionally recognized as providing a basis for a lawsuit in American courts.'" *TransUnion*, 594 U.S. at 417, 141 S.Ct. 2190 (internal quotation marks omitted). The "inquiry asks whether plaintiffs have identified a close historical or com-

mon-law analogue for their *asserted injury*," *id.* (emphasis added), not the asserted cause of action. *See Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417, 2025 WL 744032, at *1 (N.D. Cal. Mar. 7, 2025).

For both DMCA and traditional copyright infringement claims, the harm involves an injury to "an author's property right in his original work of authorship." *The Intercept Media*, 767 F.Supp.3d at 28.[6] Indeed, the DMCA was enacted by Congress "to strengthen copyright protection in the digital age," and to "combat copyright piracy," which Congress feared was "overwhelming the capacity of conventional copyright enforcement to find and enjoin unlawfully copied material." *Mango*, 970 F.3d at 170–171. The requirement under section 1202(b) that a defendant "kno[w], or, . . . hav[e] reasonable grounds to know" that their conduct "will induce, enable, facilitate, or conceal" copyright infringement "ensures that any violation of the DMCA is tied to concerns of downstream infringement." *The Intercept Media*, 767 F.Supp.3d at 28.

Accordingly, plaintiffs' allegations of harm pursuant to sections 1202(b)(1) and 1202(b)(3) are sufficiently concrete to satisfy the injury-in-fact requirement of Article III.

b. Causation

OpenAI also contends that the harm plaintiffs allege is not "fairly traceable" to

---

6. That *TransUnion* rejected the plaintiffs' allegations of harm *"absent dissemination"* does not change this outcome. There, the Supreme Court stated that *"*[t]he mere presence of an inaccuracy in an internal credit file, if it is not disclosed to a third party, causes no concrete harm.*"* *TransUnion*, 594 U.S. at 434, 141 S.Ct. 2190. However, that case involved an entirely different alleged harm (the potential dissemination of inaccurate credit informa-

tion) and an inapposite historical analogue (defamation). Because defamation requires publication of the defamatory statement, the Supreme Court concluded that those class members whose inaccurate credit information had not been publicly disseminated lacked a historical common law analogue and therefore did not suffer a *"*concrete*"* injury to as required by Article III. *See id.* at 425–26, 434, 141 S.Ct. 2190.

the purported removal of CMI from plaintiffs' works. Specifically, OpenAI contends that (1) plaintiffs have failed to establish that the alleged removal of CMI from articles in the training datasets *caused* defendants LLMs to exclude CMI from the regurgitating outputs (*see CIR*, ECF No. 100 at 18–19), and (2) the alleged harm of plaintiffs' "inability to receive speculative subscription and licensing revenue … do[es] not flow from any purported removal of CMI." (*Daily News*, ECF No. 82 at 13; *see also Times*, ECF No. 52 at 22.)

[25] The Court disagrees. To satisfy the causation requirement for Article III standing, a plaintiff must establish "that the plaintiff's injury likely was caused or likely will be caused by the defendant's conduct." *Food & Drug Admin. v. All. for Hippocratic Med.*, 602 U.S. 367, 382, 144 S.Ct. 1540, 219 L.Ed.2d 121 (2024). Put differently, "there must be a causal connection between the injury and the conduct complained of—the injury has to be fairly traceable to the challenged action of the defendant, and not the result of the independent action of some third party not before the court." *Lujan*, 504 U.S. at 560, 112 S.Ct. 2130 (alterations and internal quotation marks omitted).

[26] Here, plaintiffs allege that defendants' removal of CMI from plaintiffs' works conceals and facilitates copyright infringement, which deprives plaintiffs of licensing and subscription revenue. That harm is fairly traceable to the removal of CMI: its removal allows defendants to provide plaintiffs' works directly to end users through regurgitating outputs, while concealing that defendants infringed plaintiffs' copyrights to generate those outputs. This conduct obviates the need of end users to subscribe to plaintiffs' works or eliminates or reduces their reluctance to use defendants' products out of knowledge that doing so might constitute further in-

fringement. That the same harm could potentially occur even if defendants did *not* remove CMI from plaintiffs' works misses the point. Assuming plaintiffs would succeed on their claims that defendants removed CMI from their works—as the Court must do when determining whether plaintiffs have standing, *see City of Waukesha v. EPA*, 320 F.3d 228, 235 (D.C. Cir. 2003) (per curiam)—plaintiffs have alleged harms that are "the predictable effect" of that CMI removal. *Dep't of Com. v. New York*, 588 U.S. 752, 768, 139 S.Ct. 2551, 204 L.Ed.2d 978 (2019).

The Court finds that plaintiffs have satisfied Article III's causation requirement and have Article III standing to bring their DMCA claims.

4. Plaintiffs Have Statutory Standing To Bring Their DMCA Claims.

OpenAI's statutory standing argument also fails. OpenAI contends that even if plaintiffs suffered a DMCA violation, they are not "person[s] injured by" that violation as required by 17 U.S.C. § 1203(a) and therefore lack statutory standing.

[27] Assuming without deciding that section 1203(a) requires allegations of injury beyond a mere statutory violation, plaintiffs have satisfied this requirement. As discussed above, plaintiffs have not merely alleged that defendants violated section 1202(b); they have alleged that this violation injures them by concealing defendants' own copyright infringement, enabling and facilitating the copyright infringement of end users, diverting users from plaintiffs' websites, and causing a decline in subscription and licensing revenue.

* * *

To summarize, the Court finds that plaintiffs have established both Article III and statutory standing sufficient to enable

them to pursue their DMCA claims, because the harms they allege bear a "close relationship" to traditional copyright infringement sufficient to satisfy the injury-in-fact requirement of Article III; their alleged harms are fairly traceable to defendants' conduct; and they have alleged that they are "person[s] injured by" defendants' violation of section 1202(b), as required by section 1203(a).

B. Failure To State a Claim

On the merits, the Court finds that all three complaints fail to state a claim pursuant to section 1202(b)(1) against Microsoft. The Times also fails to state a claim pursuant to section 1202(b)(1) against OpenAI, but CIR and the *Daily News* plaintiffs have plausibly alleged that OpenAI violated section 1202(b)(1).[7] In addition, all three complaints fail to state a claim pursuant to section 1202(b)(3) against both Microsoft and OpenAI.

1. The Daily News *Plaintiffs and CIR Have Stated Claims Against OpenAI Pursuant to Section 1202(b)(1), but The Times Has Failed To Do So.*

[28] To establish a 17 U.S.C. § 1202(b)(1) violation, a plaintiff must show "(1) the existence of CMI on the allegedly infringed work, (2) the removal or alteration of that information and (3) that the removal was intentional." *Fischer v. Forrest*, 968 F.3d 216, 223 (2d Cir. 2020). Section 1202(b) also requires that defendants knew or had reasonable grounds to know that their removal of CMI would "induce, enable, facilitate, or conceal" infringement.

a. Intentional Removal of CMI

[29] OpenAI principally challenges the complaints as failing to plausibly allege that OpenAI removed CMI from the training datasets. The Times contends that because the regurgitating outputs listed in their complaints lack CMI, then *a fortiori* CMI was removed by defendants during the training process. The Court is not convinced by this argument. To the contrary, the regurgitating outputs in the *Times* complaint all contain *excerpts* of The Times' articles, not complete or substantially complete copies; it is entirely plausible that CMI remained on the articles included in the training datasets but simply did not appear in the outputs. (*See Times*, FAC Ex. J.) The *Times* complaint does not include any specific detail on how CMI was allegedly removed during the training process, and its conclusory statement that defendants' process of training their LLMs removes CMI "by design" (*id.* ¶ 187) fails to "nudge[ ] [its] claims across the line from conceivable to plausible." *Twombly*, 550 U.S. at 570, 127 S.Ct. 1955.

[30] The *Daily News* and *CIR* complaints, however, plausibly allege CMI removal during the LLM training process. The *Daily News* complaint describes OpenAI's use of the "Dragnet and Newspaper content extractors in creating the WebText dataset, which intentionally removed the [*Daily News* plaintiffs'] CMI from the [*Daily News* plaintiffs'] Works scraped from their website[s]." (*Daily News*, Compl. ¶ 161.) The *Daily News* complaint explains that Dragnet removes copyright notices "as part of the process of extracting the text content of a website," and alleges that the Newspaper content extrac-

---

7. OpenAI urges that plaintiffs' DMCA claims are time-barred to the extent they are based on the building of LLM training datasets occurring more than three years before the complaints were filed. The Court rejects these

arguments for the reasons stated in Section III. Accordingly, with respect to the surviving section 1202(b) claims, the Court denies OpenAI's motions to dismiss those claims based on the statute of limitations.

tor also "separate[s] and extract[s] the article text on the [*Daily News* plaintiffs'] webpages" while removing CMI. (*Id.*)

[31] The *CIR* complaint provides the most detail on how the Dragnet and Newspaper algorithms remove CMI during the process of assembling the dataset from which defendants trained their models. (*See CIR*, FAC ¶¶ 59–63.) CIR states that OpenAI, in developing Webtext, "used sets of algorithms called Dragnet and Newspaper to extract text from websites," and specifically alleges that "Dragnet's algorithms are designed to 'separate the main article content' from other parts of the website, including 'footers' and 'copyright notices,' and allow the extractor to make further copies only of the 'main article content.'" (*Id.* ¶¶ 59–60). In addition, Newspaper is "incapable of extracting copyright notices and footers" from the articles it scrapes from the internet, according to the complaint. (*Id.* ¶ 61.) Indeed, the *CIR* complaint alleges that OpenAI intentionally used both the Dragnet and Newspaper algorithms to "create redundancies," and that "[o]n information and belief, the OpenAI Defendants chose not to extract author and title information" when using the Newspaper algorithms "because they desired consistency with the Dragnet extractions, and Dragnet is typically unable to extract author and title information." (*Id.* ¶¶ 59, 61.) These allegations are sufficient at the pleading stage to plausibly allege that OpenAI removed CMI from CIR's works in the process of building the datasets to train their LLMs.

[32, 33] With respect to intentionality, both the *Daily News* and *CIR* complaints plausibly allege that OpenAI's CMI removal was intentional. The *Daily News* complaint explains that defendants intentionally used the Dragnet and Newspaper content extractors, which are designed to remove CMI from the works they scrape

from the internet. (*Daily News*, Compl. ¶ 161.) The *CIR* complaint alleges that "[b]ecause, by the time of its scraping, Dragnet and Newspaper were publicly known to remove author, title, copyright notices, and footers, and given that OpenAI employs highly skilled data scientists who would know how Dragnet and Newspaper work, the OpenAI Defendants intentionally and knowingly removed this copyright management information while assembling WebText." (*CIR*, FAC ¶ 64.) Plaintiffs in both the *Daily News* and *CIR* actions have satisfied their burden at the pleading stage of alleging removal of CMI, especially given this Circuit's leniency when evaluating scienter on a motion to dismiss. *See In re DDAVP Direct Purchaser Antitrust Litig.*, 585 F.3d 677, 693 (2d Cir. 2009); *Aaberg v. Francesca's Collections, Inc.*, No. 17-cv-115, 2018 WL 1583037, at *9 (S.D.N.Y. Mar. 27, 2018); *Hirsch v. CBS Broad. Inc.*, No. 17-cv-1860, 2017 WL 3393845, at *8 (S.D.N.Y. Aug. 4, 2017).

b. The Second Scienter Requirement

OpenAI also contends that plaintiffs fail to plausibly allege both (1) that the removal of CMI could enable third-party infringement, and (2) that OpenAI had knowledge of this fact. With respect to its first objection, "nothing in the statutory language [of section 1202(b)] limits its applicability to such downstream [third-party] infringement." *Mango*, 970 F.3d at 172. Section 1202(b) simply requires that "defendant know or have reason to know that distribution of copyrighted material despite the removal of CMI 'will induce, enable, facilitate, or conceal an infringement.'" *Id.* (quoting 17 U.S.C. § 1202(b)). That infringement "is not limited by actor (*i.e.*, to third parties) or by time (*i.e.*, to future conduct)," *id.*, and plaintiffs have plausibly alleged both that OpenAI's re-

moval of CMI conceals its *own* infringement, and that that removal enables and facilitates third-party infringement. *See Shihab v. Complex Media, Inc.*, No. 21-cv-6425, 2022 WL 3544149, at *5 (S.D.N.Y. Aug. 17, 2022) (finding the second scienter requirement satisfied when plaintiff plausibly alleged that the defendant knew its CMI removal would conceal "its own alleged infringement").

[34] With respect to OpenAI's second objection—that it lacked knowledge that CMI removal would induce or conceal copyright infringement—both CIR and the *Daily News* plaintiffs plausibly allege that OpenAI knew or had reason to know that its removal of CMI from plaintiffs' works would induce, enable, facilitate, or conceal copyright infringement. The complaints allege that OpenAI has publicly acknowledged both that it uses copyrighted works to train its models and that its models "are capable of distributing unlicensed copies of copyrighted works." (*See Daily News*, Compl. ¶¶ 4, 144–47, 214; *CIR*, FAC ¶¶ 83, 144.) The *CIR* complaint also alleges that OpenAI was aware of the possibility of end-user infringement through the generating of regurgitating outputs, including because of OpenAI's policy of agreeing to indemnify end users accused of infringement; its admission that its products regurgitate material in response to user prompts; and its recent adjustment to ChatGPT's settings to limit regurgitation. (*CIR*, FAC ¶¶ 83, 118.) Again, given courts' lenience in allowing issues of scienter to survive motions to dismiss when "such issues are appropriate for resolution by the trier of fact," *In re DDAVP Direct Purchaser Antitrust Litig.*, 585 F.3d at 693 (internal quotation marks omitted), the Court determines that CIR and the *Daily News* plaintiffs have plausibly alleged that OpenAI violated section 1202(b)(1).

In sum, the Court dismisses The Times's 17 U.S.C. § 1202(b)(1) claim against OpenAI and denies OpenAI's motions to dismiss the 17 U.S.C. § 1202(b)(1) claims in the *CIR* and *Daily News* complaints.

2. Plaintiffs Have Failed To State a Section 1202(b)(1) Claim Against Microsoft.

[35] Unlike the *Daily News* and *CIR* complaints' detailed allegations regarding OpenAI's removal of CMI during the process of developing its LLM training datasets, all three complaints are devoid of factual specificity to support their claims against Microsoft for violation of section 1202(b)(1). None of the allegations concerning Microsoft—including Microsoft's partnership with OpenAI to develop Copilot and Browse with Bing, and its provision of the cloud computing system which OpenAI uses to train its models—relate to any alleged removal *by Microsoft* of CMI from plaintiffs' works. To the contrary, all "the specific factual matter in the complaint related to CMI removal connects only to OpenAI." *The Intercept Media*, 767 F.Supp.3d at 32. Accordingly, the Court dismisses all three complaints' claims pursuant to section 1202(b)(1) against Microsoft.

3. Plaintiffs Have Failed To State a Section 1202(b)(3) Claim Against Defendants.

[36] To establish a section 1202(b)(3) violation, a plaintiff must show "(1) the existence of CMI in connection with a copyrighted work; and (2) that a defendant distributed works or copies of works; (3) while knowing that CMI has been removed or altered without authority of the copyright owner or the law; and (4) while knowing, or having reasonable grounds to know that such distribution will induce, enable, facilitate, or conceal an infringement." *Mango*, 970 F.3d at 171 (cleaned

up). The complaints include allegations that defendants distributed copies of plaintiffs' works both to end users and between each other. As explained below, the Court rejects both theories and concludes that plaintiffs have failed to state a claim pursuant to section 1202(b)(3) against both OpenAI and Microsoft.

a. Distribution of Copies to End Users

The Times and the *Daily News* plaintiffs contend that the "regurgitations" generated by defendants' LLMs constitute "distributions" of copies of their works. The DMCA does not define "distribution." While courts have understood "distribution" under the DMCA to require a "sale or transfer of ownership extending beyond that of a mere public display," *Wright v. Miah*, No. 22-cv-4132, 2023 WL 6219435, at *7 (E.D.N.Y. Sept. 7, 2023), and have pointed to the distinction between "distributions" and "public displays" in other parts of the Copyright Act to support this conclusion, *see e.g., FurnitureDealer.Net, Inc. v. Amazon.com, Inc.*, No. 18-cv-232, 2022 WL 891473, at *23 (D. Minn. Mar. 25, 2022) (discussing 17 U.S.C. §§ 101, 106), it is not clear whether an LLM output is a mere "public display" or something more. *Cf. id.* (reasoning that "public display does not constitute distribution, and thus is not a [DMCA] violation" and concluding that Amazon's public display of plaintiffs' product descriptions without CMI on Amazon's website did not constitute "distributions"); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1162 (9th Cir. 2007).

Assuming without deciding that the LLM regurgitations do constitute "distributions," plaintiffs have another hurdle to clear: they must show that those regurgitations constitute "works" or "copies of works." *See* 17 U.S.C. § 1202(b)(3). While the DMCA does not define "copies of works," an abundance of case law estab-

lishes that in "cases where claims of removal of CMI have been held viable, the underlying work has been substantially or entirely reproduced." *Fischer v. Forrest*, 286 F. Supp. 3d 590, 609 (S.D.N.Y. 2018) (citing *Bounce Exch., Inc. v. Zeus Enter. Ltd.*, No. 15-cv-3268, 2015 WL 8579023 (S.D.N.Y. Dec. 9, 2015)), *aff'd*, 968 F.3d 216 (2d Cir. 2020); *see also We the Protesters, Inc. v. Sinyangwe*, 724 F. Supp. 3d 281, 296 (S.D.N.Y. 2024); *cf. Doe 1 v. GitHub, Inc.*, No. 22-cv-06823, 2024 WL 235217, at *8–9 (N.D. Cal. Jan. 22, 2024).

The requirement of section 1202(b)(3) that the underlying work be "substantially or entirely reproduced" aligns with the DMCA's purpose of combatting piracy. "Fearful [about] the ease with which pirates could copy and distribute a copyrightable work in digital form," Congress sought to "strengthen copyright protection in the digital age," *Mango*, 970 F.3d at 171, by "protect[ing] the integrity of copyright management information and prohibit[ing] the removal of CMI from copyrighted works." *Shihab*, 2022 WL 3544149, at *4. Allowing DMCA claims to survive when the distributed work is not "close to identical" to the original would risk boundless DMCA liability, including liability for any person who distributes only portions of an article—e.g., various block quotes—without including CMI. *See We the Protesters*, 724 F. Supp. 3d at 297. While that person may be found liable under other causes of action, it is doubtful his reproduction of a mere portion of an article without CMI could be said to be violating the "integrity of copyright management information," *Shihab*, 2022 WL 3544149, at *4, or distributing a "work . . . [or] copies of works." 17 U.S.C. § 1202(b)(3).

[37] Reviewing the regurgitations cited in the *Times* and *Daily News* complaints, the Court concludes that the regurgitations do not constitute "substantial[ ] or

entire[ ]" reproductions of plaintiffs' works. *Fischer*, 286 F. Supp. 3d at 609. In the *Times* complaint, the regurgitations include only excerpts of the underlying articles. (*See Times*, FAC ¶¶ 99–100, 104–07, 112–22.) In addition, the excerpts were often (1) generated in response to multiple prompts, and (2) outputted in a different order than they appear in the original articles, falling far short of representing exact or substantial reproductions of the originals. With respect to the regurgitations included in Exhibit J to the *Times* complaint, those outputs also reflect only small portions of the original articles, and in fact essentially all of the regurgitations stop in the middle of a sentence. (*See, e.g.*, *Times*, FAC Ex. J at 3 (cutting off after "What's more, the company's").) These excerpts are not "work[s]" or "copies of works" as required by 17 U.S.C. § 1202(b)(3).

The regurgitations in the *Daily News* complaint contain the same flaws. The complaint includes regurgitating outputs that capture only portions of the underlying articles, as reflected in the complaint itself (*see, e.g., Daily News*, Compl. ¶¶ 98–112, 118–36), and in Exhibit J to the *Daily News* complaint.

Accordingly, because the outputs are merely excerpts of plaintiffs' works and not "copies" of those works for purposes of section 1202(b)(3), The Times and the *Daily News* plaintiffs have failed to establish that defendants "distributed" "copies" of their works in violation of section 1202(b)(3), and their section 1202(b)(3) claims against defendants are dismissed.

b. Distribution of Copies Between Defendants

Unlike the theories of distribution alleged in the *Times* and *Daily News* complaints, the *CIR* complaint alleges that OpenAI and Microsoft "distributed" CMI-less copies of CIR's works with *each other*

in violation of section 1202(b)(3). (*See CIR*, FAC ¶¶ 156–57, 168–69.) Defendants challenge these claims as failing to allege (1) distribution, (2) of works or copies of works, and (3) that defendants knew or had reason to know that such distribution would induce, enable, facilitate or conceal copyright infringement.

[38] The Court agrees that the *CIR* complaint fails to plausibly allege that OpenAI and Microsoft "distributed" CMI-less "copies" of CIR's works with each other. As set forth above, courts have understood the term "distribute" in section 1202(b)(3) to require a "sale or transfer of ownership extending beyond that of a mere public display," *Wright*, 2023 WL 6219435, at *7, and that the "copy" which is distributed must be a "substantial[ ] or entire[ ]" reproduction of the original work. *Fischer*, 286 F. Supp. 3d at 609. The *CIR* complaint asserts that Microsoft's investment in OpenAI, its provision of "the data center and bespoke supercomputing infrastructure used to train ChatGPT," and its statement on behalf of its CEO that "we have the data," give rise to the plausible inference that OpenAI and Microsoft distributed copies of CIR's works with each other. (*See CIR*, FAC ¶¶ 28–29.) As a threshold matter, neither those allegations nor any other allegation in the *CIR* complaint provide any factual support for the assertion that *Microsoft* distributed CMI-less works to OpenAI. The Court therefore dismisses CIR's section 1202(b)(3) claim against Microsoft.

[39] CIR's section 1202(b)(3) claim against OpenAI fares no better. CIR's allegations about the general business relationship between OpenAI and Microsoft fail to indicate that OpenAI "distributed" "copies" of CIR's articles to Microsoft. Nowhere in the complaint's lengthy discussion of OpenAI's process of training its

models on works scraped from the internet does CIR allege when, why, or how OpenAI would have "distributed" copies of those works to Microsoft. "[W]ithout some further factual enhancement," CIR's "formulaic recitation of the elements of a cause of action will not do." *Twombly*, 550 U.S. at 555, 557, 127 S.Ct. 1955.

CIR relies heavily on a single statement from Microsoft's CEO, stating, "we have the data." Intelligencer Staff, *Satya Nadella on Hiring the Most Powerful Man in AI*, The Intelligencer (Nov. 21, 2023), https://nymag.com/intelligencer/2023/11/on-with-kara-swisher-satya-nadella-on-hiring-sam-altman.html. However, as the court in *The Intercept Media* recently explained, that statement, considered in context, reflects Microsoft CEO Satya Nadella's "confidence in Microsoft's own AI capabilities—separate and apart from its investment in OpenAI." *The Intercept Media*, 767 F.Supp.3d at 32. Indeed, the statement was made following the Microsoft CEO's assurance that "[i]f OpenAI disappeared tomorrow, I don't want any customer of ours to be worried . . . because we have all of the rights to continue the innovation." Intelligencer Staff, *supra*. Nadella continued, "[w]e have the people, we have the compute, we have the data, we have everything." *Id.* Nadella immediately followed that statement with the clarification: "But at the same time, I'm committed to the OpenAI partnership," *id.*, suggesting that his statement "we have the data" reflected his views on Microsoft's own ability to develop AI products if OpenAI "disappeared." The interviewer then asked about the Microsoft-OpenAI partnership, and Nadella explained that Microsoft's significant investment in OpenAI "gives us significant rights," and that Microsoft

"build[s] tools" and "build[s] the infrastructure" as part of its partnership with OpenAI. *Id.*

It would constitute a significant inferential leap to conclude that the Microsoft CEO's single statement "we have the data," spoken in the context of explaining Microsoft's capacity to continue innovation and develop generative AI in a hypothetical future scenario, shows that OpenAI distributed copies of CIR's articles to Microsoft. The Court declines to make this leap and concludes that CIR's general allegations pertaining to the OpenAI and Microsoft business relationship fail to move CIR's section 1202(b)(3) claim across the line from conceivable to plausible.

* * *

For the foregoing reasons, the Court dismisses the plaintiffs' section 1202(b)(3) claims against Microsoft and OpenAI in all three actions.

VI. COMMON LAW UNFAIR COMPETITION BY MISAPPROPRIATION

The Times and the *Daily News* plaintiffs also bring claims of common law unfair competition by misappropriation. They allege that defendants engage in unfair competition by using plaintiffs' works, without authorization by plaintiffs, to train their LLMs, which in turn "produce informative text of the same general type and kind that [plaintiffs] produce[ ]." (*Times*, FAC ¶ 195; *Daily News*, Compl. ¶ 229.)[8] By producing outputs that misappropriate plaintiffs' works, defendants "directly compete with [plaintiffs'] content," free ride on plaintiffs' significant efforts to gather time-sensitive content, and harm plaintiffs through lost advertising and subscription

---

8. In this Section VI, the Court uses the term "plaintiffs" to refer to The Times and the   *Daily News* plaintiffs exclusively.

revenue, according to the complaints. (*Times*, FAC ¶¶ 194–97; *Daily News*, Compl. ¶¶ 229–31.) The *Times* complaint also alleges that defendants misappropriate its Wirecutter recommendations. (*Times*, FAC ¶¶ 193–94.) Defendants disagree and contend that plaintiffs' misappropriation claims are preempted by section 301 of the Copyright Act.

A.   Applicable Standard

Section 301 of the Copyright Act provides for "the preemption of state law claims that are interrelated with copyright claims in certain ways." *Nat'l Basketball Ass'n v. Motorola, Inc.*, 105 F.3d 841, 848 (2d Cir. 1997) ("*NBA*"). Specifically, a state law claim—such as the state-law tort of common law unfair competition by misappropriation—is preempted when that claim seeks to vindicate "legal or equitable rights that are equivalent" to any of the bundle of exclusive rights protected by copyright law. *See* 17 U.S.C. § 301(a). Section 301 sets forth a two-part test for determining whether a state law claim is preempted by the Copyright Act. First, the claim must "seek[ ] to vindicate 'legal or equitable rights that are equivalent' to one of the bundle of exclusive rights already protected by copyright law under 17 U.S.C. § 106" (the "general scope" requirement); second, the work in question must be "of the type of works protected by the Copyright Act under 17 U.S.C. §§ 102 and 103" (the "subject matter" requirement). *Barclays Cap. Inc. v. Theflyonthewall.com, Inc.*, 650 F.3d 876, 892 (2d Cir. 2011) (quoting *NBA*, 105 F.3d at 848).

[40] As a result, section 301 preempts most common law misappropriation claims involving copyrighted works. However, "certain forms of commercial misappropriation otherwise within the general scope requirement" of section 301 will survive preemption when those claims include "ex-

tra elements" instead of, or in addition to, "the acts of reproduction, performance, distribution or display . . . [that] lie within the general scope of copyright." *NBA*, 105 F.3d at 850 (citations omitted). In *NBA*, the Second Circuit held that a "hot news" misappropriation claim, originally established by the Supreme Court in *International News Service v. Associated Press*, 248 U.S. 215, 39 S.Ct. 68, 63 L.Ed. 211 (1918) ("*INS*"), contains "extra elements" and is thus not preempted by section 301 of the Copyright Act. *See NBA*, 105 F.3d at 850–53.

[41] The exception to preemption for "hot news" misappropriation claims is narrow. *See id.* at 852; *Barclays*, 650 F.3d at 897–98. To make out a "hot news" misappropriation claim, a plaintiff must at a minimum allege the following extra elements: "(i) the time-sensitive value of factual information, (ii) the free-riding by a defendant, and (iii) the threat to the very existence of the product or service provided by the plaintiff." *NBA*, 105 F.3d at 853. In *Barclays*, the Second Circuit reiterated that "[a]n indispensable element of an *INS* 'hot news' claim is free-riding by a defendant on a plaintiff's product, enabling the defendant to produce a directly competitive product for less money because it has lower costs." *Barclays*, 650 F.3d at 902 (quoting *NBA*, 105 F.3d at 854). The court described "free-riding" as the process of taking news that the plaintiff gathered and disseminated "and selling that news as though the defendant itself had gathered it." *Id.* at 903. This "unauthorized interference" with the plaintiff's "legitimate business" occurs "*precisely at the point where the profit is to be reaped*, in order to *divert a material portion of the profit* from those who have earned it to those who have not." *Id.* at 904 (quoting *INS*, 248 U.S. at 240, 39 S.Ct. 68).

In *Barclays*, several major financial institutions brought a "hot news" misappropriation claim against the defendant—a "news aggregator" and proprietor of a digital news service—alleging that the defendant was free-riding on the plaintiffs' extensive research on publicly traded companies by summarizing the plaintiffs' reports and recommendations for existing and prospective clients and publicly disseminating them on its news platform. *Id.* at 879–80. The Second Circuit concluded that the defendant was not free riding, because it was "collecting, collating and disseminating" recommendations that the plaintiff had made while "attributing the information to its source," not selling it as its own. *Id.* at 902.

B. Plaintiffs Have Failed To Plausibly Allege "Hot News" Misappropriation Claims with Respect to Their News Content and The Times's Wirecutter Recommendations.

1. Plaintiffs' News Content

The "hot news" misappropriation claims brought by The Times and the *Daily news* plaintiffs fail because they do not establish that defendants are "free-riding" on plaintiffs' news content, even if that content is time-sensitive.

[42] The prompts listed in the complaints illustrate this point. Those prompts—which query ChatGPT, Copilot, and Browse with Bing—ask the LLM to provide the text of a specific article and include the title of the article in the prompt itself. (*See, e.g., Times*, FAC ¶¶ 104, 112–21; *Daily News*, Compl. ¶¶ 98–112, 118–36.) Many prompts also provide the LLM with the name of the article's publisher. (*See e.g., Times*, FAC ¶¶ 104, 112, 121; *Daily News*, Compl. ¶¶ 98–112, 118–27.) When the prompts themselves identify the source of the article by title,

author, or publication, the complaints fail to plausibly allege non-attribution—a cornerstone of the "free riding" inquiry. *See Barclays*, 650 F.3d at 903; *Fox News Network, LLC v. TVEyes, Inc.*, 43 F. Supp. 3d 379, 399 (S.D.N.Y. 2014) (dismissing hot news misappropriation claim as preempted because "TVEyes is not passing off Fox News' content as its own"), *rev'd in part on other grounds*, 883 F.3d 169 (2d Cir. 2018). Indeed, most of the outputs restate the article's name before providing the requested information. (*See, e.g., Times*, FAC ¶¶ 104, 112, 115, 118, 121; *Daily News*, Compl. ¶¶ 104, 106, 108, 118, 121, 124.)

[43] The outputs listed in Exhibit J to the *Times* complaint and Exhibit J to the *Daily News* complaint also do not satisfy the "hot news" exception to section 301. The prompts listed in those exhibits—which query ChatGPT—fail to satisfy the time-sensitivity requirement of the "hot news" exception. Plaintiffs nowhere allege that the ChatGPT outputs in their respective Exhibits J regurgitated each article "precisely at the point where the profit [was] to be reaped." *Barclays*, 650 F.3d at 904 (quoting *INS*, 248 U.S. at 240, 39 S.Ct. 68); *see also ML Genius Holdings LLC v. Google LLC*, No. 20-3113, 2022 WL 710744, at *6 (2d Cir. Mar. 10, 2022); *Fin. Info., Inc. v. Moody's Invs. Serv., Inc.*, 808 F.2d 204, 209 (2d Cir. 1986) ("[I]mmediacy of distribution [is] necessary to sustain a 'hot' news claim."). In addition, while the prompts listed in those exhibits do not include an article's title, they do each include a direct quote from an article. Specifically, each prompt quotes a portion of an article, and the LLM outputs an additional excerpt from that same article. (*See Times*, FAC Ex. J; *Daily News*, Compl. Ex. J.) The use of direct quotes in these queries indicates that the prompter had access to at least a portion of plaintiffs'

articles at the time it queried ChatGPT, again defeating any plausible concern of non-attribution.

While the outputs listed in plaintiffs' complaints may well raise copyright concerns, they do not suffer from non-attribution, which is central to the "free riding" element of a "hot news" misappropriation claim. To the contrary, plaintiffs' common law unfair competition by misappropriation claims fall squarely within the subject matter and general scope of the Copyright Act and are therefore preempted by section 301.

2. The Times's Wirecutter Recommendations

Similarly, The Times's allegations of common law unfair competition by misappropriation with respect to its Wirecutter recommendations also are preempted by section 301 of the Copyright Act.

[44] The Times's Wirecutter recommendations do not satisfy the narrow "hot news" misappropriation exception to preemption for at least two reasons. First, the Wirecutter recommendations do not constitute time-sensitive news. Wirecutter reviews are not breaking news; they are recommendations resulting from journalists spending "tens of thousands of hours each year researching and testing products to ensure that they recommend only the best." (*Times*, FAC ¶ 128.) Indeed, The Times discusses Wirecutter in the "Reviews and Analysis" section of its complaint, separate from the complaint's "Breaking News" section, and explains that Wirecutter's research is cumulative, resulting in the "produc[tion] [of] a catalog of reviews that today covers thousands of products." (*Id.* ¶ 36.) Nothing in that description suggests time-sensitivity.

The Times suggested at oral argument that its Wirecutter recommendations are nonetheless time-sensitive based on the "inherent time sensitive nature" of Wirecutter articles containing recommendations for, e.g., "Black Friday Sales" or "Christmas presents," which are published immediately preceding those events. (*See* Tr. of Oral Arg. at 82–83.) However, nothing in the complaint alleges that the harm to The Times's Wirecutter publication is caused by defendants' misappropriation of those recommendations "precisely at the point where the profit is to be reaped." *Barclays*, 650 F.3d at 904 (citation omitted). Instead, the complaint alleges that The Times suffers harm from the removal of Wirecutter's affiliate links—from which The Times receives a commission—in defendants' LLM outputs, a harm that is not time-specific. (*Times*, FAC ¶¶ 128–29.)

[45] Second, The Times has failed to allege that its Wirecutter reviews suffer from the "free riding" of defendants. To the contrary, the outputs in the *Times* complaint that implicate Wirecutter all reference Wirecutter by name and attribute their recommendations to that publication; they do not pass off the recommendations as their own. (*See id.* ¶¶ 130–34.) Indeed, the recommendations outputted by defendants' LLMs would have little value without their attribution to Wirecutter, when the prompts specifically asked the LLMs to provide Wirecutter recommendations. *Cf. Barclays*, 650 F.3d at 903 ("It is [the defendant's] accurate attribution of the Recommendation to the creator that gives this news its value.")

It may well be true that defendants misappropriate The Times's Wirecutter recommendations, cause harm to The Times in the process, and benefit significantly from Wirecutter's recommendations without incurring the substantial costs required to generate those recommendations. But those allegations fall squarely within the general scope and subject matter of the Copyright Act, and the "extra

elements" required to support a "hot-news" exception to preemption by section 301 are not present.

Accordingly, the common law unfair competition by misappropriation claims included in plaintiffs' complaints are preempted by section 301 of the Copyright Act, and defendants' motions to dismiss those claims are granted.

VII. FEDERAL TRADEMARK DILUTION

OpenAI also moves to dismiss the federal trademark claim brought by several of the *Daily News* plaintiffs, including the New York Daily News, Chicago Tribune, Mercury News, and Denver Post (collectively, the "trademark dilution plaintiffs"). The trademark dilution plaintiffs allege they are owners of several trademarks (the "Diluted Trademarks"), which are "distinctive and 'famous marks' within the meaning of Section 43(c) of the Lanham Act, 15 U.S.C. § 1125(c) and are widely recognized by the general consuming public of the United States." (*Daily News*, Compl. ¶ 235.) They allege that defendants have used the Diluted Trademarks, without authorization, "on lower-quality and inaccurate writing," thereby "dilut[ing] the quality of the Diluted Trademarks by tarnishment, in violation of [section] 1125(c)." (*Id.* ¶¶ 246–47.) OpenAI moves to dismiss the count, contending that the complaint fails to allege that the Diluted Trademarks are "famous" under section 1125(c).

A. Applicable Standard

[46] Pursuant to 15 U.S.C. § 1125(c), "the owner of a famous mark that is distinctive" is entitled to an injunction against a party whose commercial use of that mark "is likely to cause dilution by blurring or dilution by tarnishment." 15 U.S.C. § 1125(c)(1). To state a claim for trademark dilution under section 1125(c), a plaintiff must allege: "(1) the mark is fa-

mous; (2) [the] defendant's use of the mark is made in commerce; (3) the defendant used the mark after the mark is famous; and (4) the defendant's use of the mark is likely to dilute the quality of the mark by blurring or tarnishment." *DigitAlb, Sh.a v. Setplex, LLC*, 284 F. Supp. 3d 547, 557 (S.D.N.Y. 2018) (citing *A.V.E.L.A., Inc. v. Est. of Marilyn Monroe, LLC*, 131 F. Supp. 3d 196, 211 (S.D.N.Y. 2015)). Fame is the "key ingredient" in a federal trademark dilution claim, *id.*, and section 1125(c) defines a famous mark as one that "is widely recognized by the general consuming public of the United States as a designation of source of the goods or services of the mark's owner." 15 U.S.C. § 1125(c)(2)(A). Furthermore, section 1125(c) lays out several relevant factors a court may consider in determining whether a mark is famous, including:

> (i) The duration, extent, and geographic reach of advertising and publicity of the mark, whether advertised or publicized by the owner or third parties.
> (ii) The amount, volume, and geographic extent of sales of goods or services offered under the mark.
> (iii) The extent of actual recognition of the mark.
> (iv) Whether the mark was registered under the Act of March 3, 1881, or the Act of February 20, 1905, or on the principal register.

15 U.S.C. § 1125(c)(2)(A).

[47, 48] In pleading federal trademark dilution, a complaint's "spare, conclusory allegations" that the trademark holder has "expended substantial time, effort, money, and resources advertising and promot[ing]" a trademarked product does not suffice. *CDC Newburgh Inc. v. STM Bags, LLC*, 692 F. Supp. 3d 205, 235 (S.D.N.Y. 2023). Neither does a single, conclusory allegation that a trademark is "widely recognized by the general public." *DigitAlb*,

284 F. Supp. 3d at 558. Rather, a complaint plausibly alleges that a trademark is famous when the allegations include attributes of the mark such as "nationwide recognition and respect," *e.g., Lewittes v. Cohen*, No. 03-cv-189, 2004 WL 1171261, at *6 (S.D.N.Y. 2004); continuous and pervasive use of the mark, *e.g., A.V.E.L.A.*, 131 F. Supp. 3d at 216; substantial investments in promoting and advertising the mark throughout the United States and internationally, *e.g., New York City Triathlon, LLC v. NYC Triathlon Club, Inc.*, 704 F. Supp. 2d 305, 321–22 (S.D.N.Y. 2010); "significant publicity" relating to the marks, *e.g., A.V.E.L.A.*, 131 F. Supp. 3d at 216; and that "products bearing the [plaintiff's] marks are sold throughout the United States," *id.*

B. The Trademark Dilution Plaintiffs Have Plausibly Alleged that the Diluted Trademarks Are Famous.

The trademark dilution plaintiffs allege that (i) each of their publications has been in circulation for more than 100 years (*Daily News*, Compl. ¶¶ 40–42, 45, 238–41); (ii) several of the Diluted Trademarks are federally registered (*id.* ¶ 234, Ex. I); (iii) they collectively own over 40,000 copyright registrations for works published under the Diluted Trademarks (*id.* ¶¶ 21, 22, 25–26); (iv) their publications are circulated throughout all 50 states (*id.* ¶ 243); (v) their news stories are featured by major national news outlets and have received significant attention and praise from news outlets such as CNN, MSNBC and Fox News (*id.* ¶¶ 9, 240); (vi) their publications have achieved national and international fame for their reporting of highly significant events in both U.S. and world history (*id.* ¶¶ 40–45); (vii) they invest hundreds of millions of dollars in operating their publications (*id.* ¶¶ 184, 227); (viii) millions of consumers access the trademark dilution plaintiffs' publications in print and digital format, which are circulated under the Diluted Trademarks (*id.* ¶ 242); (ix) their publications have widespread circulation across a general audience in the United States (*id.* ¶¶ 238–241); and (x) their publications have received widespread recognition for their achievements including numerous Pulitzer Prizes, which constitute the most prestigious and highly publicized national journalism award. In particular, the New York Daily News has received 11 Pulitzers (*id.* ¶ 238); The Chicago Tribune has received eight (*id.* ¶ 239); The Denver Post has received nine (*id.* ¶ 242); and The Mercury News has received two. (*Id.* ¶ 42).

[49] These allegations are a far cry from the threadbare and conclusory statements that doomed the trademark dilution cases cited by OpenAI in its motion. *Cf. DigitAlb*, 284 F. Supp. 3d at 557–58; *Glob. Brand Holdings, LLC v. Church & Dwight Co.*, No. 17-cv-6571, 2017 WL 6515419, at *5 (S.D.N.Y. Dec. 19, 2017); *Heller Inc. v. Design Within Reach, Inc.*, No. 09-cv-1909, 2009 WL 2486054, at *4 (S.D.N.Y. Aug. 14, 2009); *CDC Newburgh Inc.*, 692 F. Supp. 3d at 235.

By contrast, the allegations of the trademark dilution plaintiffs—which include detailed, factual descriptions of the nature and scope of the Diluted Trademarks' widespread circulation, recognition, achievements, and consumer subscriptions—"are sufficient to constitute a pleading that [the Diluted Trademarks are] 'famous' within the meaning of the statute." *Lewittes*, 2004 WL 1171261, at *6. The Court therefore denies OpenAI's motion to dismiss the federal trademark dilution claim in the *Daily News* action.

VIII. NEW YORK STATE TRADEMARK DILUTION

Microsoft's challenge to the state trademark dilution claim in the *Daily News*

action fares no better. That claim alleges that defendants' activities dilute the distinctiveness of the Diluted Trademarks and injure the trademark dilution plaintiffs' business reputations in violation of New York General Business Law § 360-*l*. According to the *Daily News* complaint, the dilution occurs when the outputs from defendants' LLMs falsely attribute the content of the output to the trademark dilution plaintiffs. Microsoft challenges the claim as barred by the dormant Commerce Clause of the United States Constitution.

A. Applicable Standard

[50, 51] The Commerce Clause of the U.S. Constitution, art. I, § 8, cl. 3, "not only vests Congress with the power to regulate interstate trade; the Clause also contains a further, negative command, one effectively forbidding the enforcement of certain state economic regulations even when Congress has failed to legislate on the subject." *Nat'l Pork Producers Council v. Ross*, 598 U.S. 356, 368, 143 S.Ct. 1142, 215 L.Ed.2d 336 (2023) (internal quotation marks and alterations omitted) (citing *Okla. Tax Comm'n v. Jefferson Lines, Inc.*, 514 U.S. 175, 179, 115 S.Ct. 1331, 131 L.Ed.2d 261 (1995)). This negative command has come to be known as the dormant Commerce Clause. It prevents a state from "us[ing] its laws to discriminate purposefully against out-of-state economic interests." *Id.* at 364, 143 S.Ct. 1142. Pursuant to this "antidiscrimination principle," which "lies at the very core of [the Supreme Court's] dormant Commerce Clause jurisprudence," states are prohibited from enforcing state laws motivated by "economic protectionism—that is, regulatory measures designed to benefit in-state economic interests by burdening out-of-state competitors." *Id.* at 369, 143 S.Ct. 1142 (internal quotation marks and citation omitted). By the same token, "absent discrimination, a State may exclude from its

territory, or prohibit the sale therein of any articles which, in its judgment, fairly exercised, are prejudicial to the interests of its citizens." *Id.* (internal quotation marks omitted).

[52, 53] For laws that do not purposefully discriminate against out-of-state competitors but nevertheless "incidentally burden[ ] interstate commerce," courts apply "a more permissive balancing test" set forth in *Pike v. Bruce Church, Inc.*, 397 U.S. 137, 90 S.Ct. 844, 25 L.Ed.2d 174 (1970). *Rest. L. Ctr. v. City of N.Y.*, 90 F.4th 101, 118 (2d Cir. 2024). Under the *Pike* balancing test, a state law will be struck down only "if the burden imposed on interstate commerce clearly exceeds the putative local gains." *Id.* at 118 (quoting *Town of Southold v. Town of E. Hampton*, 477 F.3d 38, 47 (2d Cir. 2007)). The *Pike* analysis primarily "serves as an important reminder that a law's practical effects may also disclose the presence of a discriminatory purpose," *Nat'l Pork*, 598 U.S. at 377, 143 S.Ct. 1142, and the *Pike* test is "most frequently deployed to detect the presence or absence of latent economic protectionism." *Id.* at 391, 143 S.Ct. 1142 (Sotomayor, J., concurring). However, a state law that appears genuinely nondiscriminatory may still violate the dormant Commerce Clause if its burdens on commerce "clearly outweigh the benefits of a state or local practice." *Dep't of Revenue of Ky. v. Davis*, 553 U.S. 328, 353, 128 S.Ct. 1801, 170 L.Ed.2d 685 (2008).

B. The New York Dilution Statute Does Not Violate the Dormant Commerce Clause.

1. N.Y. Gen. Bus. Law § 360-l Does Not Discriminate Against Out-of-State Commerce.

[54] Microsoft contends that the trademark dilution plaintiffs are not merely

seeking to enforce a state regulation that has "practical extraterritorial effects," but that they are seeking a ruling that *directly regulates* out-of-state commerce. (*See Daily News*, ECF No. 105 at 10.) That is not correct. Section 360-*l* provides grounds for injunctive relief for the dilution of a mark used on goods "sold or transported in commerce *in this state*." N.Y. Gen. Bus. Law § 360(h) (emphasis added); *id.* §§ 360(a), 360(c), 360-*l*. *Cf. Healy v. Beer Inst., Inc.*, 491 U.S. 324, 337, 109 S.Ct. 2491, 105 L.Ed.2d 275 (1989) (invalidating law requiring out-of-state beer merchants to affirm that their in-state prices were no higher than their out-of-state prices); *Brown-Forman Distillers Corp. v. N.Y. State Liquor Auth.*, 476 U.S. 573, 580, 106 S.Ct. 2080, 90 L.Ed.2d 552 (1986) (invalidating price affirming law imposed on out-of-state liquor distillers).

To the extent Microsoft posits that the "extraterritorial effects" of section 360-*l* amount to an effective regulation of wholly out-of-state commerce, the Supreme Court in *National Pork Producers Council v. Ross* rejected this theory, which would create an "almost *per se* rule forbidding enforcement of state laws that have the practical effect of controlling commerce outside the State, even when those laws do not purposely discriminate against out-of-state economic interests." 598 U.S. at 371, 143 S.Ct. 1142 (internal quotation marks and citation omitted). The Supreme Court wrote that such a theory would lead to "strange places" when, "[i]n our interconnected national marketplace, many (maybe most) state laws have the practical effect of controlling extraterritorial behavior." *Id.* at 374, 143 S.Ct. 1142 (internal quotation marks omitted). The Supreme Court confirmed that whether a state statute violates the dormant Commerce Clause depends on whether the statute discriminates against out-of-state economic interests, or in rare cases, whether a nondiscriminatory

state law substantially burdens interstate commerce in clear excess of the local benefits. *Id.* at 391, 143 S.Ct. 1142 (Sotomayor, J. concurring); *id.* at 394, 143 S.Ct. 1142 (Roberts, C.J., concurring).

Section 360-*l* "plainly does not facially discriminate against interstate commerce[,] [n]or does it harbor a discriminatory purpose." *Rest. L. Ctr.*, 90 F.4th at 106 (upholding nondiscriminatory state statute against dormant Commerce Clause challenge). Microsoft has not identified any out-of-New York-state competitor that would be disadvantaged vis-a-vis New York competitors due to section 360-*l*, nor has it articulated a discriminatory purpose against out-of-state competitors underlying the statute. Section 360-*l* does not "erect[ ] an economic barrier protecting a major local industry against competition from without the State," *Dean Milk Co. v. Madison*, 340 U.S. 349, 354, 71 S.Ct. 295, 95 L.Ed. 329 (1951); operate like "a tariff or customs duty" in order to protect New York competitors from outside competition, *W. Lynn Creamery, Inc. v. Healy*, 512 U.S. 186, 194, 114 S.Ct. 2205, 129 L.Ed.2d 157 (1994); or "deliberately rob[ ]" out-of-state competitors of "whatever competitive advantages they may possess" over in-state competitors. *Nat'l Pork*, 598 U.S. at 374, 143 S.Ct. 1142 (citation omitted). Quite the opposite, section 360-*l* fits within the myriad laws in our "interconnected national marketplace,"—including "libel laws, securities requirements, charitable registration requirements, franchise laws, tort laws," as well as "inspection laws, quarantine laws, and health laws of every description,"—that have "the practical effect of controlling extraterritorial behavior" but do not violate the Commerce Clause. *Id.* at 374–75, 143 S.Ct. 1142 (cleaned up). The New York dilution statute may "have a considerable influence on commerce outside [New York's] borders,"

id. at 375, 143 S.Ct. 1142, but Microsoft has not contended much less shown that section 360-*l* discriminates against out-of-state commerce.

2. N.Y. Gen. Bus. Law § 360-l Does Not Substantially Burden Interstate Commerce in Clear Excess of Its Local Benefits.

[55] As set forth above, the *Pike* balancing test invalidates state laws whose incidental burdens on interstate commerce clearly exceed the law's putative local gains. *Rest. L. Ctr.*, 90 F.4th at 118. In its motion to dismiss, Microsoft does not attempt to argue that section 360-*l* fails the *Pike* balancing test. For good reason: "*Pike*'s balancing and tailoring principles are most frequently deployed to detect the presence or absence of latent economic protectionism," which, as discussed, is not present in this case—nor does section 360-*l* impose a "substantial burden on interstate commerce" in clear excess of its local gains. *Nat'l Pork*, 598 U.S. at 391, 143 S.Ct. 1142 (Sotomayor, J., concurring).

Accordingly, Microsoft's motion to dismiss the trademark dilution plaintiffs' state dilution claim is denied, because Microsoft has not shown that N.Y. Gen. Bus. Law § 360-*l* discriminates against out of state commerce or that its incidental burdens on interstate commerce are in clear excess of its local benefits.

IX. DIRECT COPYRIGHT INFRINGEMENT INVOLVING ABRIDGEMENTS

Finally, in alleging direct infringement in violation of 17 U.S.C. § 501, the *CIR* complaint distinguishes between defendants' "regurgitations" of CIR's works (*CIR*, FAC ¶¶ 79–85), and their "abridgments" of those works (*id.* ¶¶ 86–98). Although CIR's definition of an "abridgment" is somewhat opaque (*see id.*), the term appears to refer to outputs that are detailed summaries of CIR articles, "often in the format of a bulleted list of main points." (*Id.* ¶ 91.)

OpenAI has moved to dismiss CIR's direct infringement claim to the extent it relates to alleged abridgments, on the grounds that the alleged abridgments, on their face, are not "substantially similar to protected expression in the articles at issue—a necessary element of a copyright infringement claim." (*CIR*, ECF No. 146 at 1.)

A. Applicable Standard

[56–58] "To establish infringement, the copyright owner must demonstrate that (1) the defendant has actually copied the plaintiff's work; *and* (2) the copying is illegal because a substantial similarity exists between the defendant's work and the protectible elements of plaintiff's." *Yurman Design, Inc. v. PAJ, Inc.*, 262 F.3d 101, 110 (2d Cir. 2001) (internal quotation marks and citation omitted); *see also Abdin v. CBS Broad. Inc.*, 971 F.3d 57, 66 (2d Cir. 2020). "The standard test in determining substantial similarity is the 'ordinary observer test': whether an average lay observer would overlook any dissimilarities between the works and would conclude that one was copied from the other." *Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc.*, 166 F.3d 65, 70 (2d Cir. 1999). When a work contains both protectible and unprotectible elements, a "more discerning" ordinary observer test applies, which asks whether there exists "substantial similarity between those elements, and only those elements, that provide copyrightability to the allegedly infringed [work]." *Boisson v. Banian, Ltd.*, 273 F.3d 262, 272 (2d Cir. 2001); *see also Abdin*, 971 F.3d at 66.

[59] The application of the "ordinary observer" test "is by no means exclusively reserved for resolution by a jury." *Peter F. Gaito Architecture, LLC v. Simone Dev.*

Corp., 602 F.3d 57, 63 (2d Cir. 2010). Indeed, "it is entirely appropriate for a district court to resolve that question as a matter of law, either because the similarity between two works concerns only noncopyrightable elements of the plaintiff's work, or because no reasonable jury, properly instructed, could find that the two works are substantially similar." *Id.* (internal quotation marks and citations omitted). Numerous district courts have resolved the substantial similarity question at the pleading stage for one or both of these reasons. *See, e.g., Montgomery v. Holland*, 408 F. Supp. 3d 353 (S.D.N.Y. 2019), *aff'd sub nom. Montgomery v. NBC Television*, 833 F. App'x 361 (2d Cir. 2020); *Nobile v. Watts*, 289 F. Supp. 3d 527 (S.D.N.Y. 2017), *aff'd*, 747 F. App'x 879 (2d Cir. 2018); *Piuggi v. Good for You Prods. LLC*, 739 F. Supp. 3d 143, 162 (S.D.N.Y. 2024).

[60, 61] Works are not substantially similar as a matter of law when their only similarity concerns general or underlying facts and ideas, which are not copyrightable. *See* 17 U.S.C. § 102(b); *see also Abdin*, 971 F.3d at 67–68. However, while facts themselves are not copyrightable, factual compilations "may possess the requisite originality" for copyright protection. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 348, 111 S.Ct. 1282, 113 L.Ed.2d 358 (1991). In such cases, where "the compilation author clothes facts with an original collocation of words, he or she may be able to claim a copyright in this written expression," and "[o]thers may copy the underlying facts from the publication, but not the precise words used to present them." *Id.* By the same token, when facts are reported "in a different arrangement, with a different sentence structure and different phrasing," the secondary work does not "purloin protected expression," and no copyright infringement has ensued. *Nihon*, 166 F.3d at 71.

B. The Abridgments Contained in the *CIR* Complaint Are Not Substantially Similar to CIR's Copyrighted Works as a Matter of Law.

[62] Exhibit 11 to the *CIR* complaint provides website links to articles that CIR alleges were unlawfully abridged by defendants in their ChatGPT and Copilot outputs. (*CIR*, FAC Ex. 11.) Examining the similarities between those outputs and the corresponding CIR articles, including the "total concept and feel, theme . . . sequence, pace, and setting," *Williams v. Crichton*, 84 F.3d 581, 588 (2d Cir. 1996), the Court concludes that the "abridgments" contained in Exhibit 11 are not substantially similar to CIR's copyrighted works as a matter of law.

The alleged abridgments are detailed summaries, usually in bullet point form, of the facts contained in CIR's articles. Those summaries—which differ in style, tone, length, and sentence structure from CIR's articles—are not "substantially similar" to CIR's copyrighted works. They present the "facts in a different arrangement"—bullet point lists or short summary paragraphs—"with a different sentence structure and different phrasing." *Nihon*, 166 F.3d at 71. In short, the abridgments in Exhibit 11 are not substantially similar, qualitatively or quantitatively, to the original CIR articles as a matter of law. The Court therefore grants OpenAI's motion to dismiss CIR's claim of direct infringement under 17 U.S.C. § 501 insofar as it relates to the "abridgments" contained in Exhibit 11.

X. CONCLUSION

For the foregoing reasons, the Court denies (1) OpenAI's motions to dismiss the direct infringement claims involving conduct occurring more than three years be-

fore the complaints were filed; (2) defendants' motions to dismiss the contributory copyright infringement claims; and (3) defendants' motions to dismiss the state and federal trademark dilution claims in the *Daily News* action.

The Court grants defendants' motions to dismiss the common law unfair competition by misappropriation claims and OpenAI's motion to dismiss the "abridgment" claims in the *CIR* action, and dismisses each of those claims with prejudice.

With respect to the DMCA claims, the Court grants (1) Microsoft's motions to dismiss the 17 U.S.C. § 1202(b)(1) claims against it in all three actions, (2) OpenAI's motion to dismiss the section 1202(b)(1) claim against it in the *Times* action, and (3) defendants' motions to dismiss the section 1202(b)(3) claims against them in all three actions, and dismisses each claim without prejudice. The Court denies OpenAI's motions to dismiss the section 1202(b)(1) claims against it in the *Daily News* and *CIR* actions.

**Richard MUELLER, II,
et al., Plaintiffs,**

v.

**DEUTSCHE BANK AKTIENGE-
SELLSCHAFT, et al., Defendants.**

24cv6225 (DLC)

United States District Court,
S.D. New York.

Signed April 4, 2025

Background: Family members and estates of people abducted and killed by a terrorist group sued a German bank and its American affiliate under the Trafficking Victims Protection Reauthorization Act, alleging that, by providing financial services to customers affiliated with terrorist group, banks knowingly benefited from participating in terrorist group's human trafficking. German bank moved to dismiss for lack of personal jurisdiction and both banks moved to dismiss for failure to state a claim.

Holdings: The District Court, Denise L. Cote, Senior District Judge, held that:

(1) plaintiffs alleged requisite connection between their claim and German bank's contacts with New York; but

(2) complaint failed to plausibly allege participation; and

(3) complaint failed to plausibly allege knowledge necessary to state a claim of beneficiary liability under the Trafficking Victims Protection Reauthorization Act.

Motion to dismiss for lack of personal jurisdiction denied.

Motion to dismiss for failure to state a claim granted.

1. Federal Courts ⚖2792

To defeat a motion to dismiss for lack of personal jurisdiction, a plaintiff must make a prima facie showing that jurisdiction exists. Fed. R. Civ. P. 12(b)(2).

2. Courts ⚖13.3(11), 13.6(1)

Courts in New York have jurisdiction over a foreign defendant for causes of action that arise out of transacting any business within the state, whether in person or through an agent. N.Y. CPLR § 302(a)(1).