

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

THE INTERCEPT MEDIA, 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.

Case No. 1:24-cv-01515-JSR

**DEFENDANT MICROSOFT CORPORATION'S MEMORANDUM
IN SUPPORT OF MOTION TO DISMISS THE COMPLAINT**

April 15, 2024

TABLE OF CONTENTS

TABLE OF AUTHORITIES ii

PRELIMINARY STATEMENT 1

ALLEGATIONS OF THE COMPLAINT..... 3

 A. Microsoft Collaborates With OpenAI To Bring GPT-Based Products To The Public..... 3

 B. The Intercept Sues Microsoft Based On Alleged Participation In Creating Or
 Distributing Datasets Used To Train GPT-Based Products..... 4

 C. The Complaint Alleges No Injury Resulting From Alleged Removal Of CMI During
 Training..... 6

ARGUMENT 7

I. THE INTERCEPT LACKS STANDING BECAUSE IT FAILS TO ALLEGE AN
ACTUAL OR THREATENED INJURY. 8

 A. Absent Dissemination, Removal Of CMI In A Non-Public Setting Inflicts No
 Concrete Harm Under Article III. 8

 B. The Intercept Fails To Establish Standing To Seek Either Damages Or
 Injunctive Relief..... 11

II. THE INTERCEPT FAILS TO STATE A DMCA § 1202 CLAIM..... 13

 A. The Complaint Does Not Plausibly Allege “Removal” Of CMI From
 Copies Of Works..... 14

 B. The Complaint Does Not (And Cannot) Plausibly Allege That *Microsoft*
 Removed CMI Or Distributed Works Lacking CMI. 18

 C. The Complaint Does Not Plausibly Allege A Likelihood That Removal
 Of CMI During Training Of An AI Tool Will Induce, Enable, Facilitate, Or
 Conceal Infringement..... 20

CONCLUSION..... 24

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Alan Ross Mach. Corp. v. Machinio Corp.</i> , No. 17-cv-3569, 2019 WL 1317664 (N.D. Ill. Mar. 22, 2019)	14
<i>Andersen v. Stability AI Ltd.</i> , No. 23-cv-00201 (N.D. Cal. Jan. 13, 2023)	2
<i>Andersen v. Stability AI Ltd.</i> , No. 23-cv-00201-WHO, 2023 WL 7132064 (N.D. Cal. Oct. 30, 2023)	17
<i>Ashcroft v. Iqbal</i> , 556 U.S. 662 (2009)	7, 16
<i>Burress v. Freedom Mortg. Corp.</i> , No. 20-cv-15242-NLH-AMD, 2022 WL 2916070 (D.N.J. July 22, 2022)	9
<i>Carter v. Helmsley-Spear, Inc.</i> , 71 F.3d 77 (2d Cir. 1995)	10
<i>Cartoon Network LP, LLLP v. CSC Holdings, Inc.</i> , 536 F.3d 121 (2d Cir. 2008)	19
<i>Clapper v. Amnesty Int’l USA</i> , 568 U.S. 398 (2013)	12
<i>Cole v. John Wiley & Sons, Inc.</i> , No. 11-cv-2090-DF, 2012 WL 3133520 (S.D.N.Y. Aug. 1, 2012)	17, 18
<i>Dastar Corp. v. Twentieth Century Fox Film Corp.</i> , 539 U.S. 23 (2003)	10
<i>Dinerstein v. Google, LLC</i> , 73 F.4th 502 (7th Cir. 2023)	9
<i>Doe 1 v. GitHub, Inc.</i> , 672 F. Supp. 3d 837 (N.D. Cal. 2023)	13
<i>Doe 1 v. GitHub, Inc.</i> , No. 22-cv-06823 (N.D. Cal. Nov. 3, 2022)	1
<i>Doe 1 v. GitHub, Inc.</i> , No. 22-cv-06823-JST, 2024 WL 235217 (N.D. Cal. Jan. 22, 2024)	13, 16, 17

Dolls Kill, Inc. v. Zoetop Bus. Co.,
 No. 22-cv-01463-RGK-MAA, 2022 WL 16961477 (C.D. Cal. Aug. 25, 2022).....17

Fashion Nova, LLC v. Blush Mark, Inc.,
 No. 22-cv-6127, 2023 WL 4307646 (C.D. Cal. June 30, 2023).....11

First Nationwide Bank v. Gelt Funding Corp.,
 27 F.3d 763 (2d Cir. 1994).....16

Free Speech Sys., LLC v. Menzel,
 390 F. Supp. 3d 1162 (N.D. Cal. 2019)17, 22

Gilliam v. Am. Broad. Cos.,
 538 F.2d 14 (2d Cir. 1976).....10

Harrington v. Pinterest, Inc.,
 No. 20-cv-05290-EJD, 2022 WL 4348460 (N.D. Cal. Sept. 19, 2022).....20, 23

Kadrey v. Meta Platforms, Inc.,
 No. 23-cv-03417 (N.D. Cal. July 7, 2023).....2

Kadrey v. Meta Platforms, Inc.,
 No. 23-cv-03417-VC, 2023 WL 8039640 (N.D. Cal. Nov. 20, 2023)22

Kent v. Universal Studios, Inc.,
 No. 08-cv-2703-GAF, 2008 WL 11338293 (C.D. Cal. Aug. 15, 2008).....10

Kirk Kara Corp. v. W. Stone & Metal Corp.,
 No. 20-cv-1931-DMG, 2020 WL 5991503 (C.D. Cal. Aug. 14, 2020).....17

Lujan v. Defenders of Wildlife,
 504 U.S. 555 (1992).....8, 11

Mango v. BuzzFeed, Inc.,
 970 F.3d 167 (2d Cir. 2020).....20

Melendez v. Sirius XM Radio, Inc.,
 50 F.4th 294 (2d Cir. 2022)17

Mills v. Netflix, Inc.,
 No. 19-cv-7618-CBM, 2020 WL 548558 (C.D. Cal. Feb. 3, 2020).....23

Philpot v. Altnet Media, Inc.,
 No. 18-cv-04479-TSH, 2018 WL 6267876 (N.D. Cal. Nov. 30, 2018)23

Roberts v. BroadwayHD LLC,
 518 F. Supp. 3d 719 (S.D.N.Y. 2021).....11, 21

Steel v. Bongiovi,
784 F. Supp. 2d 94 (D. Mass. 2011)14

Stevens v. Corelogic, Inc.,
899 F.3d 666 (9th Cir. 2018)21, 22, 23

Suid v. Newsweek Mag.,
503 F. Supp. 146 (D.D.C. 1980).....10

Tilford v. Jones,
No. 05-cv-2989, 2006 WL 2612752 (S.D. Tex. Sept. 11, 2006).....10

TransUnion LLC v. Ramirez,
594 U.S. 413 (2021)..... *passim*

Tremblay v. OpenAI, Inc.,
No. 23-cv-03223 (N.D. Cal. June 28, 2023).....1, 2, 16

Tremblay v. OpenAI, Inc.,
No. 23-cv-3223-AMO, 2024 WL 557720 (N.D. Cal. Feb. 12, 2024)16, 20, 23

VanderKodde v. Elliott,
No. 17-cv-203-PLM, 2023 WL 9898588 (W.D. Mich. Apr. 27, 2023)9

Statutes

Copyright Act, 17 U.S.C. §§ 101 et seq.

§ 106A.....10

§ 1202..... *passim*

§ 1202(b)..... *passim*

§ 1202(b)(1)4, 13, 14

§ 1202(b)(3)4, 13, 14, 19

§ 1203(a) 2, 14

Fair Credit Reporting Act8

Fair Debt Collection Practices Act9

Federal Rules of Civil Procedure

Rule 121

Rule 12(b)(1).....7, 8

Rule 12(b)(6).....	7, 13
Truth in Lending Act	9
Other Authorities	
3 Patry on Copyright § 9:5.50 (2024)	19
<i>Landing Page</i> , Common Crawl, https://www.commoncrawl.org (last visited April 13, 2024)	18
S. Rep. 105-190 (1998).....	11

PRELIMINARY STATEMENT

Large Language Models, or LLMs, and the applications that use them, like ChatGPT, have awed the public with their capabilities and potential. These generative artificial intelligence tools are created by feeding large amounts of text through a machine learning algorithm. The algorithm learns statistical relationships among constituent “tokens” in the text (words, punctuation, and the like). This allows an LLM to generate brand new natural language content in response to user prompts.

The Intercept is not the first to challenge this technology through claims under the Digital Millennium Copyright Act’s provision concerning removal of copyright management information (CMI), such as book titles or authors’ names. 17 U.S.C. § 1202(b). Other plaintiffs have alleged that either in the process of training the LLM, or as a result of the LLM’s generated outputs, the AI tools “remove” CMI from copies of copyrighted works in a fashion that may facilitate infringement of those works. The other LLM-based § 1202(b) claims have all been dismissed under Rule 12. The same result is appropriate here. Indeed, the § 1202(b) claims brought by The Intercept in this case are the skimpiest of the lot of these challenges.

At the outset, The Intercept’s suit should be dismissed for lack of standing. The Intercept has chosen to bring § 1202(b) claims solely based on alleged removal of CMI during the *training* of an LLM—it does not allege that a copy of its works lacking CMI has ever been publicly disseminated. Doubtless it has done so because every plaintiff that has tried to bring a § 1202(b) claim based on an LLM’s *outputs* has run into the reality that LLMs simply do not emit the sorts of full copies of identical copyrighted works necessary to sustain a CMI claim.¹ This gambit to

¹ See, e.g., *Doe 1 v. GitHub, Inc.*, No. 22-cv-06823 (N.D. Cal. Nov. 3, 2022); *Tremblay v. OpenAI, Inc.*, No. 23-cv-03223 (N.D. Cal. June 28, 2023); *Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417 (N.D. Cal. July 7, 2023); *Andersen v. Stability AI Ltd.*, No. 23-cv-00201 (N.D. Cal. Jan. 13, 2023).

evade the flaws of other cases does not work, because even if The Intercept had plausibly alleged removal of CMI in the non-public setting of training—which it has not—that conduct results in no concrete injury. The mere presence of a copy of a work without a plaintiff’s CMI, absent dissemination, is precisely the sort of vaporous technical harm that the Supreme Court has found insufficient to confer standing. *See TransUnion LLC v. Ramirez*, 594 U.S. 413, 423 (2021). It also is not enough to sustain a claim under the DMCA; since the purpose of CMI is to inform the public that a work is copyrighted, alleged removal of CMI out of public view does not result in an “injured party” under 17 U.S.C. § 1203(a). Because The Intercept alleges no more, it lacks both standing, *infra* § I, and any right to sue, *infra* § II.

Even if The Intercept alleged sufficient injury to bring this suit, it fails to state a claim. Its entire Complaint is little more than generalizations designed to insinuate that CMI must have been removed from its works during LLM training, without any factual matter rendering that inference plausible. The Complaint asserts that “thousands” of The Intercept’s unidentified works of journalism were contained in “training sets” that Microsoft and OpenAI used to train AI tools. But it provides no basis for this claim of mass copying—it does not identify any specific copyrighted works or even assert copyright infringement claims. It similarly provides no basis for the claim that when Defendants created training sets, they removed The Intercept’s CMI. The Complaint just baldly asserts that ChatGPT’s outputs “generally” do not contain CMI, and that they would surely contain CMI if the LLMs were trained on data containing CMI, as though this means that anything in the training set *must* have been stripped of CMI. Once again, the Complaint offers nothing—no technical details on how ChatGPT works, description of real-world user behavior, or example of ChatGPT’s operation—to lend plausibility to the claim. The allegation that Defendants removed CMI from The Intercept’s works is pure hypothesis of the

sort courts have uniformly rejected when other plaintiffs have challenged LLMs with § 1202 claims. *Infra* § II.A.

Beyond this fundamental defect, the Complaint fails to support its speculation that *Microsoft* participated in any training of OpenAI’s LLM, and therefore any removal of CMI or distribution of works lacking CMI. *Infra* § II.B. Finally, The Intercept fails to plausibly establish that the alleged removal of CMI from its works during training would have any likelihood of facilitating or concealing piracy, as required for a § 1202(b) claim. *Infra* § II.C. The Court should dismiss the Complaint in its entirety.

ALLEGATIONS OF THE COMPLAINT

A. Microsoft Collaborates With OpenAI To Bring GPT-Based Products To The Public.

Over the past several years, researchers and engineers have made dramatic advances in the field of AI. The LLM is one of those advances. LLMs extract the elements of language by “ingest[ing] massive amounts” of information based on large bodies of data called “training sets.” Compl. ¶ 4. This process is called “training.” Once trained, an LLM can generate natural-language responses to user prompts, “mimick[ing] how humans write and speak.” Compl. ¶ 4. Defendant OpenAI’s “GPT” models are among the most well-known LLMs.

According to the Complaint, “Defendants are the companies responsible for the creation and development of ... ChatGPT AI products,” Compl. ¶ 5—products built on the GPT models that “provide[] responses to questions or other prompts” from users, Compl. ¶ 34. Though short on details, the Complaint asserts that Microsoft “invested billions of dollars” into OpenAI and “provide[d] the data center and supercomputing infrastructure used to train ChatGPT” and to “host[] ChatGPT training sets.” Compl. ¶¶ 20-22. The Complaint mentions “Microsoft’s own Bing Copilot,” Compl. ¶ 44, and “Bing AI products,” Compl. ¶ 67, but offers no further allegations concerning the development, operation, or use of either.

B. The Intercept Sues Microsoft Based On Alleged Participation In Creating Or Distributing Datasets Used To Train GPT-Based Products.

The Intercept is a “news organization” that provides a platform for “human journalists” who produce works based on “in-depth investigations.” Compl. ¶¶ 6-8. It alleges that Defendants, in allegedly creating training datasets, “intentionally removed author, title, copyright notice, and terms of use information from Plaintiff’s copyrighted works.” Compl. ¶¶ 43-44. The Complaint does not identify any specific copyrighted work allegedly included in a training dataset. Nor does the Complaint identify any output of an LLM allegedly containing all or a portion of the contents of any such copyrighted work. Nevertheless, The Intercept brings two causes of action against each Defendant under 17 U.S.C. § 1202: one under § 1202(b)(1) for “intentionally remov[ing] or alter[ing] any copyright management information,” *id.*, when Defendants allegedly created datasets, Compl. ¶¶ 51-63 (Count I); Compl. ¶¶ 66-75 (Count III); the other under § 1202(b)(3) for “distribut[ing] ... copies of works ... knowing that copyright management information has been removed or altered,” *id.*, when they allegedly shared those datasets (or works in them) with each other, Compl. ¶¶ 64-65 (Count II); Compl. ¶¶ 76-77 (Count IV). The Intercept brings no copyright infringement claim.

The Complaint offers little factual substance concerning what works it claims are in the training set for the products at issue; what Defendants allegedly did with respect to these training sets or the training process; and what tangible effects have resulted from this alleged conduct. It instead relies on chains of inferences allegedly drawn from unspecified “publicly available information,” Compl. ¶ 29, that is “available on the internet” but not cited in the Complaint, Compl. ¶ 38, plus undescribed “consultations with a data scientist.” Compl. ¶ 29.

According to the Complaint, “[b]ased on the publicly available information ..., thousands of Plaintiff’s copyrighted works were included in Defendants’ training sets.” Compl. ¶ 42. The

Intercept acknowledges that it does not know what “content” was in the actual training sets Defendants used, Compl. ¶ 29, and it never identifies a single Intercept work that it claims was included. Instead, it asserts that “versions of ChatGPT ... were trained using at least the following training sets: WebText, WebText2, and Common Crawl.” Compl. ¶ 30. The Complaint alleges that “[v]arious sources have recreated approximations of” these datasets, and that “thousands” of unidentified Intercept works are contained in these anonymous approximations. Compl. ¶¶ 38, 42. The Complaint does not cite any of these so-called sources. (The Complaint also does not contain allegations about the datasets used to train Microsoft’s “Bing Copilot” or “Bing AI.”)

The Complaint gets more elliptical from there. Though it repeatedly asserts that CMI was removed from The Intercept’s works, Compl. ¶¶ 43-44, it offers no substantiation of this. It relies instead on generalities. The Complaint asserts, apparently based on the aforementioned consultations with a “data scientist,” Compl. ¶ 29, that “[i]f ChatGPT was trained on works of journalism that included the original author, title, copyright notice, and terms of use information, ChatGPT would have learned to communicate that information,” Compl. ¶ 39. The Complaint provides neither technical detail nor illustration substantiating this. It then asserts that “[w]hen ChatGPT provides responses to users, it generally does not provide” this information. Compl. ¶ 40. Again, the Complaint offers no factual matter supporting this assertion, nor even an example of it happening with any work. It offers only the conclusory syllogism that ChatGPT “generally” does not output CMI; that it would if it were trained on CMI; and therefore that *everything* in the training set *must have* been stripped of CMI.

Thinnest of all are allegations that suggest that Microsoft engaged in the above conduct. The Complaint lacks plausible allegations that *Microsoft* removed CMI from The Intercept’s

works in creating datasets, or that *Microsoft* distributed The Intercept’s works knowing that the CMI had been removed from those works. For the former, the Complaint offers only bare assertion: “WebText and WebText2 were created by the OpenAI Defendants. Common Crawl originated elsewhere, but was adapted and utilized by Defendants.... *Upon information and belief, ... Defendant Microsoft [has] created Common Crawl datasets, as opposed to copying a dataset already created by someone other than Defendants.*” Compl. ¶ 31 (emphasis added). The Intercept does not (and cannot) offer factual substance supporting the italicized assertion.

As for the alleged “distribution” of copies of works with CMI removed, the Complaint’s theory seems to be that because of “Microsoft’s provision of database and computing resources to the OpenAI Defendants, Defendant Microsoft has shared copies of Plaintiff’s works from which [CMI] has been removed.” Compl. ¶ 45. The Complaint alleges no more to support that theory.

C. The Complaint Alleges No Injury Resulting From Alleged Removal Of CMI During Training.

As far as the Complaint discloses, The Intercept has suffered no concrete harm as a result of the alleged removal of CMI from its “thousands” of works. Compl. ¶ 42. To be sure, it floats a conclusory theory of how removal of CMI *could* cause a copyright holder harm. It posits that if “ChatGPT provid[es] responses to ChatGPT users that incorporate[] or regurgitate[] material verbatim from copyrighted works,” those users might “further distribute the results,” but “would be less likely to [do so] if they were made aware of the author, title, copyright, and terms of use information.” Compl. ¶¶ 47-49. This is an attempt to satisfy § 1202(b)’s scienter requirement, which demands that the defendant intended, knew, or should have known that removal of CMI would facilitate or conceal infringement. The Complaint provides no examples of its theory occurring in the real world; no support for its assessment of user behavior; no allegations

suggesting that user distribution of ChatGPT outputs would be infringing; and nothing to quantify this supposed risk.

The Complaint also points to no instance in which a copy of an Intercept copyrighted work has ever been emitted by ChatGPT, Bing Copilot, or a Bing AI product. It offers no allegations about the likelihood that any such copy of any such work will be contained in a response by one of those products. And it alleges no other form of harm resulting from the mere removal of CMI from some unidentified work during the process of training an LLM.

ARGUMENT

To survive a motion to dismiss under Rule 12(b)(1) brought based on a lack of Article III standing, “a plaintiff must show (i) that [it] suffered an injury in fact that is concrete, particularized, and actual or imminent; (ii) that the injury was likely caused by the defendant; and (iii) that the injury would likely be redressed by judicial relief.” *TransUnion LLC v. Ramirez*, 594 U.S. 413, 423 (2021). Whether an injury is concrete for purposes of Article III depends on whether the “plaintiff[] ha[s] identified a close historical or common-law analogue for their asserted injury”; that inquiry cannot depend “on contemporary, evolving beliefs about what kinds of suits should be heard in federal courts.” *Id.* at 424-25.

To survive a motion to dismiss under Rule 12(b)(6), “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 (2009) (quotation marks and citation omitted). If a claim lacks a cognizable theory, lacks sufficient facts to plausibly support that theory, or advances a theory that is foreclosed as a matter of law, dismissal is appropriate. “Threadbare recitals of the elements of a cause of action, supported by mere conclusory statements,” cannot state a claim. *Id.*

I. THE INTERCEPT LACKS STANDING BECAUSE IT FAILS TO ALLEGE AN ACTUAL OR THREATENED INJURY.

The Complaint fails at the threshold under Rule 12(b)(1) because The Intercept has not alleged injury-in-fact and therefore lacks standing. Even if The Intercept’s allegations that Microsoft removed CMI from copies of The Intercept’s works were plausible (they are not, *see infra* §§ II.A., II.B.), the Complaint does not allege that these CMI-less copies were disseminated to the public. The mere removal of CMI in a non-public setting from a copy that never sees the light of day causes no harm. Because that is all The Intercept alleges, it lacks standing.

A. Absent Dissemination, Removal Of CMI In A Non-Public Setting Inflicts No Concrete Harm Under Article III.

To establish standing, plaintiffs must “demonstrate ... that they suffered a concrete harm”—“real, and not abstract.” *TransUnion*, 594 U.S. at 417, 424 (cleaned up). “Congress may ‘elevate’ harms that ‘exist’ in the real world before Congress recognized them to actionable legal status, [but] 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.” *Id.* at 426 (cleaned up). Only those harms with “a close historical or common-law analogue for the[] asserted injury” can confer standing. *Id.* at 424. And it is not enough for a plaintiff to allege an injury to *someone’s* “cognizable interest”—a plaintiff “seeking review” must itself be “among the injured.” *Lujan v. Defenders of Wildlife*, 504 U.S. 555, 563 (1992) (cleaned up).

The Supreme Court applied these principles in *TransUnion v. Ramirez*. The plaintiffs in *TransUnion* “alleged violations of the Fair Credit Reporting Act” for failing “to follow reasonable procedures to ensure the accuracy of credit files so that the files would not include [Office of Foreign Assets Control] alerts labeling the plaintiffs as potential terrorists.” 594 U.S. at 430. For a subset of these plaintiffs, defendant TransUnion did indeed maintain records with “misleading OFAC alerts,” but had never “provide[d] those plaintiffs’ credit information to any

potential creditors.” *Id.* at 433. All “assumed” that TransUnion had committed a statutory violation. *Id.* at 432 n.5. The question was whether these plaintiffs had suffered a concrete injury merely based on the non-public maintenance of false information about them.

The Supreme Court held that they had not. The Court explained 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.” *Id.* at 434. Reasoning that “[p]ublication” is essential to a claim for defamation—the closest common-law analogue to a claim based on false credit information—the Court held that “no historical or common-law analog” supports an injury based on “the mere existence of inaccurate information.” *Id.* (quotation marks and citation omitted). Nor could those plaintiffs gain standing by restyling their injury as a “material risk of future harm” that inaccurate information might be disseminated. *Id.* at 435. Harm must “materialize” to support a “suit for damages.” *Id.* at 436. And even to maintain a suit for injunctive relief, a plaintiff must show that the risk of harm is “imminent and substantial.” *Id.* at 435. Courts have applied *TransUnion*’s framework to find lack of concrete injury in various contexts.²

The Intercept’s claims are remarkably similar to those that failed in *TransUnion*. It invokes a statute of recent vintage that for the first time endows a type of information—CMI—with legal import, then creates a novel legal violation for its removal from a copy. The Intercept at most claims nothing but a bare technical violation of that statute. It points to no instance in which copies of its works lacking CMI were ever disseminated, nor substantiates the idea that

² *Dinerstein v. Google, LLC*, 73 F.4th 502, 512-16 (7th Cir. 2023) (finding lack of Article III injury for state-law privacy claim based on use of anonymized patient records to create AI healthcare product); *VanderKodde v. Elliott*, No. 17-cv-203-PLM, 2023 WL 9898588, at *5 (W.D. Mich. Apr. 27, 2023) (no Article III injury underlying alleged violations of the Fair Debt Collection Practices Act); *Burress v. Freedom Mortg. Corp.*, No. 20-cv-15242-NLH-AMD, 2022 WL 2916070, at *4 (D.N.J. July 22, 2022) (same, for alleged violations of the Truth in Lending Act).

there is imminent and substantial risk of non-attribution of its copyrighted works to the public. And indeed, unlike the false and defamatory information at issue in *TransUnion*, The Intercept alleges only the *absence* of information about its *ownership* of a work. The only harm The Intercept asks us to deduce (and yet still fails to allege) is thus the private *non-attribution* of its *ownership* of unidentified Intercept works to The Intercept—which is to say, no harm at all.

Nothing in the common law remotely suggests otherwise. In general, the mere lack of attribution—even public non-attribution—“is not a cognizable cause of action under ... common law.” *Tilford v. Jones*, No. 05-cv-2989, 2006 WL 2612752, at *4 (S.D. Tex. Sept. 11, 2006); *see also Suid v. Newsweek Mag.*, 503 F. Supp. 146, 149 (D.D.C. 1980) (noting that the court had “been unable to locate[] any case recognizing a common-law action for failure to attribute”). American courts have at times found rights of attribution in “the tort of unfair competition” or “state unfair competition laws.” *Gilliam v. Am. Broad. Cos.*, 538 F.2d 14, 24 (2d Cir. 1976); *but see Kent v. Universal Studios, Inc.*, No. 08-cv-2703-GAF, 2008 WL 11338293, at *4 (C.D. Cal. Aug. 15, 2008) (finding that *Gilliam* has been overruled by *Dastar Corp. v. Twentieth Century Fox Film Corp.*, 539 U.S. 23 (2003)). But even if such a theory were viable, it would be based on “the author’s personal right to prevent the presentation of his work *to the public* in a distorted form” or to prevent “a false impression of ... origin.” *Gilliam*, 538 F.2d at 24 (emphasis added); *see also Carter v. Helmsley-Spear, Inc.*, 71 F.3d 77, 81 (2d Cir. 1995) (recognizing that the “civil law” right of attribution is “*personal to the artist*”) (emphasis added); *cf.* 17 U.S.C. § 106A (narrow right of action for human “authors” of visual works to “claim authorship”).

Not only is there no common-law tradition suggesting that private non-attribution of ownership information constitutes an injury—Congress also would not have viewed it as an actionable statutory injury when it enacted § 1202(b). The aim of the DMCA was to “discourage

piracy,” S. Rep. 105-190 at 11 n.18 (1998), “assist in tracking and monitoring uses of *copyrighted works*,” *id.* at 16 (emphasis added), and foster an “efficient Internet marketplace,” *id.* “[T]he purpose of CMI is to provide *the public* with notice that a work is copyrighted.” *Fashion Nova, LLC v. Blush Mark, Inc.*, No. 22-cv-6127, 2023 WL 4307646, at *5 (C.D. Cal. June 30, 2023); *Roberts v. BroadwayHD LLC*, 518 F. Supp. 3d 719, 737 (S.D.N.Y. 2021) (same). The reason for protecting CMI is not some freestanding interest in having attribution information attached to a copy of a work, but to prevent harms from *public dissemination* of that work without CMI. Removal of CMI without dissemination of a copy of a work therefore does not result in even the sort of injury Congress had in mind when it enacted the DMCA.

B. The Intercept Fails To Establish Standing To Seek Either Damages Or Injunctive Relief.

The Complaint’s failure to allege injury leaves The Intercept without standing to pursue either damages or injunctive relief.

With respect to damages, The Intercept would need to plausibly allege that some injury has “materialize[d],” *supra* 9—that is, that a GPT-based product has actually disseminated a copy of its works without CMI. *TransUnion*, 594 U.S. at 436. The Complaint alleges generally that “[a]t least some of the time, ChatGPT provides or has provided responses to users that regurgitate verbatim or nearly verbatim copyright-protected works of journalism.” Compl. ¶ 35. It provides no examples of any such works, and certainly not Intercept works. That could never suffice—otherwise the plaintiffs in *TransUnion* could have simply said “[a]t least some of the time” false OFAC information is disseminated. A plaintiff must itself be “among the injured” to claim standing. *Lujan*, 504 U.S. at 563.

Nor does the Complaint establish injury by asserting that “Microsoft had reason to know that” ChatGPT and Bing AI products would “provide responses to users that incorporated

material from Plaintiff’s copyright-protected works or regurgitated copyright-protected works verbatim.” Compl. ¶ 71. The purpose of this conclusory (and thus weightless) allegation is to satisfy § 1202(b)’s scienter requirement, which Plaintiff fails to do. *Infra* §§ II.B., II.C. But in any event, the Complaint never plausibly alleges that any such output actually happened, and therefore alleges no injury that can be redressed by this Court.

As for injunctive relief, The Intercept seeks “[a]n injunction requiring Defendants to remove all copies of Plaintiff’s copyrighted works from which [CMI] was removed from their training sets and any other repositories.” Compl. at 14, ¶ (ii). To the extent the only injury for which The Intercept seeks injunctive relief is the mere presence of works in Defendants’ training sets, that is not a redressable injury for reasons already explained.

It may be that The Intercept means to predicate standing for injunctive relief on the purported risk that Microsoft may at some point disseminate a copy of one of The Intercept’s works without CMI. If so, the Complaint is woefully deficient. It contains no allegations concerning how any AI product even works—for example, the length of its outputs or the manner in which it responds to prompts. And even if it did, the likelihood that any GPT-based product would emit an entire copy of one of The Intercept’s works depends heavily on *user* behavior—another topic about which the Complaint says nothing. Courts should be “reluctan[t] to endorse standing theories that rest on speculation about the decisions of independent actors.” *Clapper v. Amnesty Int’l USA*, 568 U.S. 398, 414 (2013). The Intercept’s nascent theory is far too speculative to confer standing based on a risk of future harm. *Id.* at 401 (a future injury is speculative if it is not “certainly impending”).

A recent case involving a challenge to a GPT-based product, *Doe 1 v. GitHub, Inc.*, is instructive here. 672 F. Supp. 3d 837 (N.D. Cal. 2023). The plaintiffs in *Doe 1* alleged

violations of § 1202 by claiming that an AI code-completion tool emitted suggestions that matched code contained in public GitHub repositories, without reproducing CMI. *Id.* at 850. On a first motion to dismiss, the court held that the plaintiffs lacked standing to pursue damages because they could not allege that Copilot had ever emitted suggestions matching *their* code. *Id.* at 850-51. But it allowed injunctive relief claims to proceed based on allegations concerning the operation of the AI tool at issue and the purported likelihood that plaintiffs' code would match that tool's output, *id.*—allegations that are entirely lacking here.

But matters did not end there. The plaintiffs amended their complaint to try to plausibly allege standing for damages, but in doing so had to reveal meaningful detail about the actual operation of the AI tool at issue and the nature of its output. Once they did so, they wound up “plead[ing] themselves out of their Section 1202(b)(1) and 1202(b)(3) claims” entirely by showing that the AI tool would likely *never* disseminate a full copy of their works at all, and that their artfully pled allegations concerning the risk of a future Section 1202 violation were in fact empty. *Doe 1 v. GitHub, Inc.*, No. 22-cv-06823-JST, 2024 WL 235217, at *8 (N.D. Cal. Jan. 22, 2024). *Doe 1* thus stands for the proposition that bare allegations that unidentified works were included *in a training set* for an AI tool are insufficient to allege that the tool has or will cause injury. And it strongly suggests that if The Intercept were required to allege factual detail about how the AI tools at issue work, it would be unable to state any claim at all.

II. THE INTERCEPT FAILS TO STATE A DMCA § 1202 CLAIM.

The Intercept's claims should also be dismissed under Rule 12(b)(6) because the Complaint fails to state a claim under either § 1202(b)(1) or (b)(3) of the DMCA. Compl. ¶¶ 66-77. At the threshold, for reasons explained above, The Intercept cannot sue under the DMCA because it has not alleged facts rendering it a “person injured” under the statute. 17 U.S.C.

§ 1203(a); *see Steel v. Bongiovi*, 784 F. Supp. 2d 94, 97-98 (D. Mass. 2011); *Alan Ross Mach. Corp. v. Machinio Corp.*, No. 17-cv-3569, 2019 WL 1317664, at *4 (N.D. Ill. Mar. 22, 2019).

The Intercept’s § 1202 claims against Microsoft also fail on their own terms. To state a claim under § 1202(b)(1), a plaintiff must allege that the defendant “intentionally remov[ed] or alter[ed]” CMI from a copy of the plaintiff’s work. 17 U.S.C. § 1202(b)(1). To state a claim under § 1202(b)(3), a plaintiff must plausibly allege that the defendant “distribute[d]” copies of the plaintiff’s works “knowing that [CMI] has been removed or altered without authority of the copyright owner[.]” And for both claims, a plaintiff must further establish that the defendant removed or distributed CMI “knowing, or ... having reasonable grounds to know, that it will induce, enable, facilitate, or conceal an infringement of any right under this title.” 17 U.S.C. § 1202(b).

The Complaint’s attempt to plead these claims is defective in three independent respects. First, The Intercept fails to plausibly allege removal of CMI from its works. *Infra* § II.A. Second, the Intercept fails to plausibly allege that *Microsoft* removed CMI from its works or distributed its works without CMI. *Infra* § II.B. And third, The Intercept fails to plausibly allege that any removal of CMI or distribution of works would have an objective likelihood of facilitating infringement or thwarting efforts to police it. *Infra* § II.C.

A. The Complaint Does Not Plausibly Allege “Removal” Of CMI From Copies Of Works.

To state a claim under § 1202(b)(1) or (b)(3), a plaintiff must plausibly allege the removal (or alteration) of CMI from copies of the plaintiff’s works. 17 U.S.C. § 1202(b)(1) (“intentionally remove or alter...”); *id.* § 1202(b)(3) (“distribute ... copies of works ... knowing that [CMI] has been removed or altered”). The Complaint fails to do so.

The Complaint does not identify or describe any copyrighted work from which CMI was allegedly removed. It provides no examples of works allegedly contained in training sets used to train any of the products at issue. It does not claim to have ever seen a copy of its works shorn of CMI. And it offers no instance in which one of the AI products at issue emitted text that matched one of its works, whether accompanied by CMI or not. Indeed, the Complaint fails to provide a single instance of this happening with *any* work.

The Intercept instead strings together a daisy-chain of generalities. It first claims to have learned “[b]ased on ... publicly available information” that “thousands of Plaintiff’s copyrighted works were included in Defendants’ training sets.” Compl. ¶ 42. The Intercept does not identify *what* public information and *who* supposedly provided this insight, offering the Court no ability to evaluate whether the inference it draws is a reasonable one. Even accepting this completely unsubstantiated assertion as true, the mere presence of a work in a training set would not, of course, establish that CMI was removed from that work. So The Intercept goes on to allege, again without supporting factual matter, that “[w]hen ChatGPT provides responses to users, it generally does not provide” CMI. Compl. ¶ 40. According to the Complaint (again with no supporting factual allegations), “[i]f ChatGPT *was* trained on works ... that included” CMI, it “would have learned to communicate that information.” Compl. ¶ 39 (emphasis added). The notion is that anything that *comes out* of ChatGPT without CMI must have had that CMI removed during training. And then comes the true Olympian leap: That this all somehow means that *everything* ever included in the training set—even things that have never been “provide[d]” as “responses to users,” Compl. ¶ 39—must have been stripped of CMI.

This daisy-chain of hypotheticals is insufficient for a viable claim for several reasons. To begin with, it does not logically follow from the general inference that if CMI was allegedly

removed from some works then CMI was removed from *all* works. This Court need not accept as true The Intercept’s “unwarranted deductions of fact.” *First Nationwide Bank v. Gelt Funding Corp.*, 27 F.3d 763, 771 (2d Cir. 1994) (quotation marks omitted). The Intercept must offer reasons to believe that CMI was removed from *its own* works to “nudge[] [its] claims ... across the line from conceivable to plausible.” *Iqbal*, 556 U.S. at 680-81 (quotation marks omitted). Another plaintiff’s failure to do so is why the court in *Tremblay v. OpenAI, Inc.*, a different case involving ChatGPT, dismissed § 1202(b) claims that were based on the same generalized suggestion that “[b]y design, the training process does not preserve any CMI,” Tremblay Compl. ¶ 64, No. 23-cv-3223-AMO (June 28, 2023), ECF No. 1; *Tremblay*, No. 23-cv-3223-AMO, 2024 WL 557720, at *4-5 (N.D. Cal. Feb. 12, 2024) (holding that “knowing removal of CMI from ... books during the training process” does not violate § 1202(b)).

The premises that make up The Intercept’s logic are also not supported by “sufficient factual matter” to be accepted. *Iqbal*, 556 U.S. at 677. In particular, the Complaint offers no basis even for the “general[]”—that is, non-Intercept-specific—assertion that ChatGPT “generally” does not provide CMI, nor that it always would do so if CMI were not removed during training. Indeed, the Complaint raises more questions than it answers on these fronts. For example, the Complaint says that “[a]t least some of the time, ChatGPT provides or has provided responses to users that mimic *significant amounts* of material from copyright-protected works ... without providing any [CMI].” Compl. ¶ 36 (emphasis added). But if, hypothetically, ChatGPT provided only an excerpt of part of an article in response to a user query, it is hard to see why it would be expected to provide the author’s byline from the top or how that could ever state a CMI claim, which requires the removal of CMI from identical works. *See Doe 1*, 2024 WL 235217, at *8-9 (dismissing § 1202(b) claim for failure to meet the pleading requirement

that CMI was removed from identical copies of works); *Dolls Kill, Inc. v. Zoetop Bus. Co.*, No. 22-cv-01463-RGK-MAA, 2022 WL 16961477, at *4 (C.D. Cal. Aug. 25, 2022) (same); *Kirk Kara Corp. v. W. Stone & Metal Corp.*, No. 20-cv-1931-DMG, 2020 WL 5991503, at *6 (C.D. Cal. Aug. 14, 2020) (same). In any event, it is The Intercept’s job to offer a plausible, not just hypothetical, explanation of why its premises hold. As the Second Circuit has explained, “conclusory assertions ... of an impermissible use [and] hypotheticals in support of those conclusions” cannot “support a plausible ... claim.” *Melendez v. Sirius XM Radio, Inc.*, 50 F.4th 294, 307 (2d Cir. 2022).

And last, even if this Court were to accept The Intercept’s basic theory of *generalized* CMI removal, that still would not be enough, because it does not identify its *own* works from which CMI has been removed. A plaintiff must allege facts demonstrating “what the removed or altered CMI was.” *Free Speech Sys., LLC v. Menzel*, 390 F. Supp. 3d 1162, 1175 (N.D. Cal. 2019). In *Free Speech Systems*, the court dismissed a complaint for failing to describe “which [photographs] had [CMI] removed or what the removed CMI was.” *Id.* at 1166. The Intercept does not even attempt to satisfy this standard, offering up only general allegations about its “works of journalism” that were “published on the internet,” Compl. ¶ 32; nods to so-called “thousands of [its] copyrighted works [that] were included in Defendants’ training sets,” without identifying a single one, Compl. ¶ 42; and threadbare recitations of various forms of CMI, e.g., Compl. ¶ 32. The Intercept’s failure to identify particular works and the specific CMI that was allegedly removed from those particular works warrants dismissal. *See Andersen v. Stability AI Ltd.*, No. 23-cv-00201-WHO, 2023 WL 7132064, at *11 (N.D. Cal. Oct. 30, 2023) (rejecting § 1202 claim where plaintiffs offered conclusory allegations that failed to identify “particular” CMI that was removed from the claimed works); *see also Cole v. John Wiley & Sons, Inc.*, No.

11-cv-2090-DF, 2012 WL 3133520, at *12 (S.D.N.Y. Aug. 1, 2012) (dismissing complaint based on “vague and expansive allegations regarding which works are the subject of Plaintiff’s claims”).

B. The Complaint Does Not (And Cannot) Plausibly Allege That *Microsoft* Removed CMI Or Distributed Works Lacking CMI.

Even if the Complaint plausibly alleged that CMI was removed from The Intercept’s works, it does not plausibly allege (and cannot allege) that *Microsoft* either (a) removed that CMI or (b) distributed those works without that CMI. Beyond its formulaic recitations of claim elements, *see* Compl. ¶¶ 67-77, the Complaint deploys three tactics to try to insinuate that *Microsoft* engaged in such conduct. All three fall far short.

First, the Complaint alleges that the “Common Crawl” dataset was used to train ChatGPT, and then baldly asserts that, “Upon information and belief, ... Defendant *Microsoft* [has] created Common Crawl datasets, as opposed to copying a dataset already created by someone other than Defendants.” Compl. ¶ 31. The Complaint offers no basis for this information and belief, nor details of any inquiry it conducted to form such a belief. The allegation is also illogical. The Complaint acknowledges that “Common Crawl originated elsewhere.” Compl. ¶ 31; *see generally* *Landing Page*, Common Crawl, <https://www.commoncrawl.org> (last visited April 13, 2024) (“Common Crawl maintains a free, open repository of web crawl data that can be used by anyone.”). The Intercept does not explain what it could possibly mean for *Microsoft* to create a dataset that originates elsewhere. And in any event, The Intercept does not allege that this mythical *Microsoft*-created version of Common Crawl was used to train OpenAI’s ChatGPT product—especially when the Complaint also alleges that *OpenAI* “created Common Crawl datasets.” Compl. ¶ 31.

Second, the Complaint vaguely alludes to “the relationship between Defendant Microsoft and the OpenAI Defendants,” Compl. ¶ 45, seemingly to insinuate that anything that occurred in OpenAI’s development of ChatGPT, Microsoft must have done as well simply by virtue of their relationship. The Complaint alleges nothing about who from Microsoft was allegedly involved in the OpenAI ChatGPT training process, or what those actors allegedly did.

Third, the Complaint alleges that “Microsoft created and hosted the data centers used to develop ChatGPT and information regarding Microsoft’s own Bing Copilot,” and that “Microsoft’s provision of database and computing resources” means that “Microsoft has shared copies of Plaintiff’s works from which [CMI] had been removed, with the OpenAI Defendants.” Compl. ¶¶ 44-45. The contention appears to be that by providing OpenAI with general-purpose computing resources that it used to host datasets, Microsoft “distributed” those datasets within the meaning of § 1202(b)(3).

This conflicts with Second Circuit principles in the copyright context holding that provision of technology that someone *else* uses to engage in some conduct in connection with a work—copying, distribution, display, and so forth—is not the same as engaging in that conduct oneself. This “volitional conduct” requirement focuses on what “causes” a copy or distribution “to be made.” *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*, 536 F.3d 121, 131 (2d Cir. 2008); 3 Patry on Copyright § 9:5.50 (2024) (explaining that volitional conduct requires that a “defendant must have chosen to engage in the conduct that is adjudged to be infringing” and noting that it is “lacking where a defendant’s system is merely used to create a copy by a third party”) (second quote cleaned up). Where a defendant does nothing more than “design[], house[], and maintain[] a system” that someone else uses to make a copy or initiate a distribution, the defendant has not engaged in that conduct. *Cartoon Network*, 536 F.3d at 131.

This principle forecloses The Intercept’s argument that Microsoft, by hosting works, actively and knowingly distributes those works. Were it otherwise, every cloud storage provider or DVR manufacturer would be automatically liable for distributing works without CMI any time its equipment automatically copied or transmitted a copy made by someone else. Nothing in § 1202(b)’s text, structure, or purpose supports such sweeping liability.

Because The Intercept fails to plausibly allege that Microsoft itself removed CMI or knowingly distributed works lacking it, this Court should do as others have done and dismiss The Intercept’s § 1202(b) claims. *See, e.g., Harrington v. Pinterest, Inc.*, No. 20-cv-05290-EJD, 2022 WL 4348460, at *4 (N.D. Cal. Sept. 19, 2022) (dismissing § 1202 claim where complaint lacked plausible allegations that defendant “intentionally removed CMI from [plaintiff’s] works or that [defendant] distributed his works knowing that CMI had been removed”) (capitalization altered); *Tremblay*, 2024 WL 557720, at *4 (same).

C. The Complaint Does Not Plausibly Allege A Likelihood That Removal Of CMI During Training Of An AI Tool Will Induce, Enable, Facilitate, Or Conceal Infringement.

Last, the claims should be dismissed for the independent reason that the Complaint fails to plausibly allege that any removal of CMI would be likely to “induce, enable, facilitate, or conceal an infringement.” 17 U.S.C. § 1202(b). Because it is an anti-piracy statute, § 1202 is careful not to impose liability on a defendant whose removal of CMI has only the incidental effect of aiding or concealing copyright infringement. It enacts this limitation through a stringent “double-scienter requirement.” *Mango v. BuzzFeed, Inc.*, 970 F.3d 167, 172 (2d Cir. 2020). The first half of the requirement demands that the defendant’s conduct—removal of CMI, alteration of CMI, and so forth—be intentional (for subsection (b)(1)) or knowing (for subsection (b)(3)). *See id.* The second half, at issue here, requires that the “defendant know or have reason to know” that its conduct “will induce, enable, facilitate, or conceal an infringement.” *Id.*

Though ultimately a scienter requirement, courts have recognized that it is founded on an objective component. Unless a plaintiff can point to direct factual allegations concerning the defendant's intent or belief that removal of CMI will facilitate infringement a plaintiff must at a minimum show that resulting infringement is objectively "likely." *Stevens v. Corelogic, Inc.*, 899 F.3d 666, 675 (9th Cir. 2018). Otherwise there is no basis for inferring that a defendant "was aware or had reasonable grounds to be aware of the probable future impact of its actions." *Id.* at 674.

The Intercept does not even try to allege that Microsoft removed CMI intending to cause infringement by anyone else. It also fails to allege a likelihood of such infringement. The Complaint's speculative theory first alleges that inclusion of a work in an AI tool's training set "would result in ChatGPT providing responses to ChatGPT users that incorporated or regurgitated material verbatim from copyrighted works," but without including CMI. Compl. ¶ 47. If this happens, the Complaint continues, then "users of ChatGPT would further distribute the results." Compl. ¶ 48. But, the Complaint claims, they "would be less likely to distribute ChatGPT responses if they were made aware of [CMI]." Compl. ¶ 49. As before, none of this is substantiated by factual allegations.

This theory is woefully deficient. To start, courts have looked skeptically upon arguments that removal of CMI in a non-public setting is likely to facilitate or conceal infringement. The point of CMI, and thus § 1202's protections, is to "inform *the public* that something is copyrighted." *Roberts*, 518 F. Supp. 3d at 737 (quotation marks omitted and emphasis added). As discussed, training happens in private and there is no allegation that training data is ever disseminated outside OpenAI. As such, it is highly *unlikely* that CMI removal during training could "induce, enable, facilitate, or conceal an infringement" by users of

Defendants’ tools, or anyone else for that matter. 17 U.S.C. § 1202(b). Without factual allegations establishing, rather than nakedly asserting, an actual likelihood that copies of the plaintiff’s works will be exposed publicly, a plaintiff cannot show that the absence of CMI would make any difference. *See Free Speech Sys.*, 390 F. Supp. at 1175. *Kadrey v. Meta Platforms, Inc.* applied the same rule in the context of the sort of LLM at issue in this case, No. 23-cv-03417-VC, 2023 WL 8039640, at *2 (N.D. Cal. Nov. 20, 2023) (dismissing complaint for failure to allege “facts to support the allegation that [the LLM] ever distributed the plaintiffs’ books”).

Nor is it sufficient to merely assert generally that end users will infringe by further distributing ChatGPT responses, or to speculate that they would be deterred from doing so if CMI were present. That is just a lightly-dressed recitation of the claim element—that lack of CMI will “induce, enable, [or] facilitate ... an infringement,” 17 U.S.C. § 1202(b). What *The Intercept* must allege is factual matter making it plausible that removal of CMI from *The Intercept’s* works will result in end-user infringement of *The Intercept’s* works. But it does not identify any of the works at issue; point to an instance of any work being output by ChatGPT; cite an example of someone using its works in an infringing way (versus, say, making a fair use); show that CMI has a meaningful deterrent effect on end users; or otherwise lend substance to its bare theory.

The Ninth Circuit’s decision in *Stevens* shows why this is insufficient. 899 F.3d at 675. *Stevens* involved the defendant’s removal of CMI metadata from photographs as part of technological “[d]ownsampling”—i.e., compression. *Id.* at 671. Plaintiff photographers suggested that this somehow would aid infringement. But they “ha[d] not ... averred that they ha[d] ever used CMI metadata to prevent or detect copyright infringement.” *Id.* at 675. Nor could they establish that the defendant was aware of a third-party “pattern of conduct” or

“established modus operandi” that might give rise to an inference that the defendant was “aware of the probable future impact of its actions” with respect to CMI. *Id.* at 674 (quotation marks omitted). Without such specifics, the plaintiffs had “simply identif[ied] a general possibility that exists whenever CMI is removed,” not a likelihood that it will occur. *Id.* at 673-75.

The same is true here. The Intercept has not alleged that Microsoft intended to cause infringement. It also has alleged no objectively plausible basis for its bald assertions concerning user behavior. Courts routinely dismiss complaints on this basis. *See Tremblay*, 2024 WL 557720, at *4-5 (dismissing CMI claim based on bare allegation that if CMI is removed, “ChatGPT users will not know if any output is infringing”); *Mills v. Netflix, Inc.*, No. 19-cv-7618-CBM, 2020 WL 548558, at *3 (C.D. Cal. Feb. 3, 2020) (dismissing § 1202 claim because there were no plausible allegations linking CMI removal to future infringement or substantiating that defendants were aware that CMI removal “would cause future infringement”); *Pinterest, Inc.*, 2022 WL 4348460, at *5 (dismissing § 1202 claim where complaint lacked plausible allegations that defendant had knowledge about how plaintiff used CMI to police infringement); *Philpot v. Alternet Media, Inc.*, No. 18-cv-04479-TSH, 2018 WL 6267876, at *5 (N.D. Cal. Nov. 30, 2018) (dismissing § 1202 claim where plaintiff “fail[ed] to plead any facts showing that [defendant] had the required mental state”). The Court should do the same here.

CONCLUSION

The Court should grant the motion to dismiss in its entirety.

Dated: April 15, 2024

Orrick, Herrington & Sutcliffe LLP

San Francisco, CA

By: *s/ Annette L. Hurst*

Annette L. Hurst (admitted *pro hac vice*)
The Orrick Building
405 Howard Street
San Francisco, CA 94105-2669
Telephone: (415) 773-5700
ahurst@orrick.com

Lisa T. Simpson
51 West 52nd Street
New York, NY 10019
Telephone: (212) 506-3767
lsimpson@orrick.com

Christopher J. Cariello
51 West 52nd Street
New York, NY 10019
Telephone: (212) 506-3778
ccariello@orrick.com

*Attorneys for Defendant
Microsoft Corporation*