

United States District Court  
Northern District of California

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA

J. DOE 1, et al.,  
Plaintiffs,  
v.  
GITHUB, INC., et al.,  
Defendants.

Case No. 22-cv-06823-JST

**ORDER GRANTING IN PART AND  
DENYING IN PART MOTIONS TO  
DISMISS**

Re: ECF Nos. 50, 53

Before the Court are motions to dismiss filed by Defendants GitHub, Inc. and Microsoft Corporation, ECF No. 50; and Defendants OpenAI, Inc., OpenAI, L.P., OpenAI GP, L.L.C., OpenAI Startup Fund GP I, L.L.C., OpenAI Startup Fund I, L.P., and OpenAI Startup Fund Management, LLC (collectively “OpenAI Defendants”), ECF No. 53. Court will grant the motions in part and deny them in part.

**I. BACKGROUND**

Plaintiffs are software developers who challenge Defendants’ development and operation of Copilot and Codex, two artificial intelligence-based coding tools.<sup>1</sup> For the purposes of the present motions, the Court accepts as true the following facts in the operative complaint.<sup>2</sup>

GitHub, which was acquired by Microsoft in 2018, is the largest internet hosting service for software projects stored in Git, a widely used open-source version control system for managing

---

<sup>1</sup> Plaintiffs define “artificial intelligence” (“AI”) as “a computer program that algorithmically simulates human reasoning or inference, often using statistical methods.” Compl. ¶ 2.

<sup>2</sup> By stipulation and order, ECF No. 47, the operative complaint in the consolidated case is that filed in *Doe 3 et al. v. GitHub, Inc. et al.*, No. 22-cv-7074-JST, ECF No. 1. All references to the complaint in this order refer to the operative complaint in the consolidated case.

1 software source code. Using GitHub permits software developers or programmers to collaborate  
 2 on projects stored in repositories. Repositories may be private or public; anyone can view and  
 3 access code stored in public repositories.

4 All code uploaded to GitHub is subject to the GitHub Terms of Service, which provide that  
 5 users retain ownership of any content they upload to GitHub, but grant GitHub the “right to store,  
 6 archive, parse, and display [the content], and make incidental copies, as necessary to provide the  
 7 Service, including improving the Service over time.” No. 22-cv-7074-JST, ECF No. 1-2 at 27.  
 8 This “includes the right to do things like copy [the code] to our database and make backups; show  
 9 it to you and other users; parse it into a search index or otherwise analyze it on our servers; [and]  
 10 share it with other users.” *Id.* at 27-28. Further, the Terms of Service provide that users who set  
 11 their repositories to be viewed publicly “grant each User of GitHub a nonexclusive, worldwide  
 12 license to use, display, and perform [the content] through the GitHub Service and to reproduce  
 13 [the content] solely on GitHub as permitted through GitHub’s functionality.” *Id.* at 28.

14 Without AI-based assistance, programmers generally write code “both by originating code  
 15 from the writer’s own knowledge of how to write code as well as by finding pre-written portions  
 16 of code that—under the terms of the applicable license—may be incorporated into the coding  
 17 project.” Compl. ¶ 78. Plaintiffs have each published licensed materials in which they hold a  
 18 copyright interest to public repositories on GitHub. When creating a new repository, a GitHub  
 19 user may “select[] one of thirteen licenses from a dropdown menu to apply to the contents of that  
 20 repository.” *Id.* ¶ 34 n.4. Two of the suggested licenses waive copyrights and related rights. The  
 21 remaining eleven suggested licenses<sup>3</sup> require that any derivative work or copy of the licensed work  
 22 include attribution to the owner, inclusion of a copyright notice, and inclusion of the license terms.  
 23 Each Plaintiff published code to a public repository on GitHub under one of the eleven suggested  
 24

---

25 <sup>3</sup> These eleven licenses are (1) Apache License 2.0; (2) GNU General Public License version 3  
 26 (“GPL-3.0”); (3) MIT License; (4) The 2-Clause BSD License (“BSD 2”); (5) The 3-Clause BSD  
 27 License (“BSD 3”); (6) Boost Software License; (7) Eclipse Public License 2.0; (8) GNU Affero  
 28 General Public License version 3 (“AGPL-3.0”); GNU General Public License version 2 (“GPL  
 2”); (10) GNU Lesser General Public License version 2.1 (“LGPL-2.1”); and (11) Mozilla Public  
 License 2.0. Compl. ¶ 34 n.4.

1 licenses that include these three requirements.

2 In June 2021, GitHub and OpenAI released Copilot, an AI-based program that can “assist  
3 software coders by providing or filling in blocks of code using AI.” *Id.* ¶ 8. In August 2021,  
4 OpenAI released Codex, an AI-based program “which converts natural language into code and is  
5 integrated into Copilot.” *Id.* ¶ 9. Codex is integrated into Copilot: “GitHub Copilot uses the  
6 OpenAI Codex to suggest code and entire functions in real-time, right from your editor.” *Id.* ¶ 47  
7 (quoting GitHub website). GitHub users pay \$10 per month or \$100 per year for access to  
8 Copilot. *Id.* ¶ 8.

9 Codex and Copilot employ machine learning, “a subset of AI in which the behavior of the  
10 program is derived from studying a corpus of material called training data.” *Id.* ¶ 2. Using this  
11 data, “through a complex probabilistic process, [these programs] predict what the most likely  
12 solution to a given prompt a user would input is.” *Id.* ¶ 79. Codex and Copilot were trained on  
13 “billions of lines” of publicly available code, including code from public GitHub repositories. *Id.*  
14 ¶¶ 82-83.

15 Despite the fact that much of the code in public GitHub repositories is subject to open-  
16 source licenses which restrict its use, *id.* ¶ 20, Codex and Copilot “were not programmed to treat  
17 attribution, copyright notices, and license terms as legally essential,” *id.* ¶ 80. Copilot reproduces  
18 licensed code used in training data as output with missing or incorrect attribution, copyright  
19 notices, and license terms. *Id.* ¶¶ 56, 71, 74, 87-89. This violates the open-source licenses of  
20 “tens of thousands—possibly millions—of software developers.” *Id.* ¶ 140. Plaintiffs additionally  
21 allege that Defendants improperly used Plaintiffs’ “sensitive personal data” by incorporating the  
22 data into Copilot and therefore selling and exposing it to third parties. *Id.* ¶¶ 225-39.

23 Plaintiffs filed multiple cases against Defendants, which were subsequently consolidated.  
24 ECF No. 47. Plaintiffs, on behalf of themselves and two putative classes,<sup>4</sup> plead twelve counts

25 \_\_\_\_\_  
26 <sup>4</sup> The “Injunctive Relief Class” and “Damages Class” are each defined as: “All persons or entities  
27 domiciled in the United States that[] (1) owned an interest in at least one [U.S.] copyright in any  
28 work; (2) offered that work under one of GitHub’s Suggested Licenses; and (3) stored Licensed  
Materials in any public GitHub repositories at any time between January 1, 2015 and the present  
(the ‘Class Period’).” Compl. ¶ 34 (footnote omitted).

1 against Defendants: (1) violation of the Digital Millennium Copyright Act (“DMCA”), 17 U.S.C.  
2 §§ 1201-05; (2) common law breach of open-source licenses; (3) common law tortious  
3 interference in a contractual relationship; (4) common law fraud; (5) false designation of origin in  
4 violation of the Lanham Act, 15 U.S.C. § 1125; (6) unjust enrichment in violation of Cal. Bus. &  
5 Prof. Code §§ 17200, *et seq.*, and the common law; (7) unfair competition in violation of the  
6 Lanham Act, 15 U.S.C. § 1125; Cal. Bus. & Prof. Code §§ 17200, *et seq.*, and the common law;  
7 (8) breach of contract for violation of the GitHub Privacy Policy and Terms of Service; (9)  
8 violation of the California Consumer Privacy Act (“CCPA”); (10) common law negligence; (11)  
9 common law civil conspiracy; and (12) declaratory relief under 28 U.S.C. § 2201(a) and Cal. Code  
10 Civ. Proc. § 1060.<sup>5</sup> Defendants now move to dismiss the complaint. ECF Nos. 50, 53.

## 11 **II. JURISDICTION**

12 The Court has jurisdiction over Plaintiffs’ federal claims under 28 U.S.C. § 1331 and  
13 supplemental jurisdiction over Plaintiffs’ state law claims under 28 U.S.C. § 1367.

## 14 **III. LEGAL STANDARD**

### 15 **A. Rule 12(b)(1)**

16 “Article III of the Constitution confines the federal judicial power to the resolution of  
17 ‘Cases’ and ‘Controversies.’” *TransUnion LLC v. Ramirez*, 141 S. Ct. 2190, 2203 (2021). “For  
18 there to be a case or controversy under Article III, the plaintiff must have a ‘personal stake’ in the  
19 case—in other words, standing.” *Id.* (quoting *Raines v. Byrd*, 521 U.S. 811, 819 (1997)). A  
20 defendant may attack a plaintiff’s assertion of jurisdiction by moving to dismiss under Rule  
21 12(b)(1) of the Federal Rules of Civil Procedure. *Cetacean Cmty. v. Bush*, 386 F.3d 1169, 1174

---

22  
23 <sup>5</sup> While Plaintiffs plead several common law tort claims, they do not identify the state law which  
24 applies to each claim. “[D]ue to variances among state laws, failure to allege which state law  
25 governs a common law claim is grounds for dismissal.” *In re Nexus 6P Prods. Liab. Litig.*, 293 F.  
26 Supp. 3d 888, 933 (N.D. Cal. 2018) (quoting *Romero v. Flowers Bakeries, LLC*, No. 14-cv-05189-  
27 BLF, 2016 WL 469370, at \*12 (N.D. Cal. Feb. 8, 2016)). With respect to the state law claims in  
28 any future amended complaint, Plaintiffs shall identify the state under whose law the claim is  
brought. When claims which share a legal theory are brought under multiple laws (for example, a  
federal statute, a state statute, and common law, such as Plaintiffs’ claim for unfair competition),  
Plaintiffs shall state each claim as a separate count. For the purposes of deciding the present  
motion to dismiss, the Court interprets the complaint as asserting the common law claims of  
Counts 2, 3, 4, 6, 7, 10, and 11 under California law.

1 (9th Cir. 2004); *see also Maya v. Centex Corp.*, 658 F.3d 1060, 1067 (9th Cir. 2011) (“[L]ack of  
2 Article III standing requires dismissal for lack of subject matter jurisdiction under Federal Rule of  
3 Civil Procedure 12(b)(1).”).

4 “A Rule 12(b)(1) jurisdictional attack may be facial or factual.” *Safe Air for Everyone v.*  
5 *Meyer*, 373 F.3d 1035, 1039 (9th Cir. 2004). “In a facial attack, the challenger asserts that the  
6 allegations contained in a complaint are insufficient on their face to invoke federal jurisdiction.”  
7 *Id.* Where, as here, a defendant makes a facial attack, the court assumes that the complaint’s  
8 allegations are true and draws all reasonable inferences in the plaintiff’s favor. *Wolfe v.*  
9 *Strankman*, 392 F.3d 358, 362 (9th Cir. 2004).

#### 10 **B. Rule 12(b)(6)**

11 “Dismissal under [Federal Rule of Civil Procedure] 12(b)(6) is appropriate only where the  
12 complaint lacks a cognizable legal theory or sufficient facts to support a cognizable legal theory.”  
13 *Mendondo v. Centinela Hosp. Med. Ctr.*, 521 F.3d 1097, 1104 (9th Cir. 2008). A complaint must  
14 contain “a short and plain statement of the claim showing that the pleader is entitled to relief.”  
15 Fed. R. Civ. P. 8(a)(2). Facts pleaded by a plaintiff “must be enough to raise a right to relief  
16 above the speculative level.” *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 555 (2007).

17 “To survive a motion to dismiss, a complaint must contain sufficient factual matter,  
18 accepted as true, to ‘state a claim to relief that is plausible on its face.’” *Ashcroft v. Iqbal*, 556 U.S.  
19 662, 678 (2009) (quoting *Twombly*, 550 U.S. at 570)). “A claim has facial plausibility when the  
20 plaintiff pleads factual content that allows the court to draw the reasonable inference that the  
21 defendant is liable for the misconduct alleged.” *Id.* In determining whether a plaintiff has met this  
22 plausibility standard, the Court must “accept all factual allegations in the complaint as true and  
23 construe the pleadings in the light most favorable to the nonmoving party.” *Knievel v. ESPN*, 393  
24 F.3d 1068, 1072 (9th Cir. 2005).

#### 25 **C. Leave to Amend**

26 Leave to amend a complaint “shall be freely given when justice so requires.” Fed. R. Civ.  
27 P. 15(a)(2). The decision of whether to grant leave to amend is “within the discretion of the  
28 district court, which may deny leave due to ‘undue delay, bad faith or dilatory motive on the part

1 of the movant, repeated failure to cure deficiencies by amendments previously allowed, undue  
 2 prejudice to the opposing party by virtue of allowance of the amendment, and futility of  
 3 amendment.” *Leadsinger, Inc. v. BMG Music Pub.*, 512 F.3d 522, 532 (9th Cir. 2008) (quoting  
 4 *Foman v. Davis*, 371 U.S. 178, 182 (1962)).

#### 5 **IV. DISCUSSION**

##### 6 **A. Article III Standing**

7 Defendants argue that Plaintiffs lack Article III standing to assert their claims and that the  
 8 Court therefore lacks subject-matter jurisdiction over this action.<sup>6</sup>

9 “[T]o establish standing, a plaintiff must show (i) that he suffered an injury in fact that is  
 10 concrete, particularized, and actual or imminent; (ii) that the injury was likely caused by the  
 11 defendant, and (iii) that the injury would likely be redressed by judicial relief.” *TransUnion*, 141  
 12 S. Ct. at 2203. “The party invoking federal jurisdiction bears the burden of establishing these  
 13 elements.” *Lujan v. Defenders of Wildlife*, 504 U.S. 555, 561 (1992). “Where, as here, a case is at  
 14 the pleading stage, the plaintiff must ‘clearly . . . allege facts demonstrating’ each element.”  
 15 *Spokeo, Inc. v. Robins*, 578 U.S. 330, 338 (2016) (quoting *Warth v. Seldin*, 422 U.S. 490, 518  
 16 (1975)).

17 “A plaintiff must demonstrate constitutional standing separately for each form of relief  
 18 requested.” *Davidson v. Kimberly-Clark*, 889 F.3d 956, 969 (9th Cir. 2018). In a putative class  
 19 action, named plaintiffs “must allege and show that they personally have been injured, not that  
 20 injury has been suffered by other, unidentified members of the class to which they belong and  
 21 which they purport to represent.” *Warth*, 422 U.S. at 502.

22 Plaintiffs advance two main theories of harm. One theory is that Defendants have shared,

---

23  
 24 <sup>6</sup> Both parties impermissibly incorporate substantive material by reference across their briefs.  
 25 “The incorporation of substantive material by reference is not sanctioned by the federal rules.”  
 26 *Swanson v. U.S. Forest Serv.*, 87 F.3d 339, 345 (9th Cir. 1996). In the future, all arguments the  
 27 parties seek to make must be raised in the briefing on a particular motion. *See Woolfson v. Conn*  
 28 *Appliances, Inc.*, No. 21-cv-07833-MMC, 2022 WL 3139522, at \*6 (N.D. Cal. Aug. 5, 2022)  
 (declining to consider substantive arguments purportedly incorporated by reference); *Williams v.*  
*Cnty. of Alameda*, 26 F. Supp. 3d 925, 947 (N.D. Cal. 2014) (“[T]he Court will not consider the  
 arguments that Plaintiff improperly seeks to incorporate by reference. This Court only considers  
 arguments that are specifically and distinctively raised by the parties in their briefs.”).

1 sold, and exposed and will continue to share, sell, and expose Plaintiffs’ personal information,  
2 harming Plaintiffs’ rights to privacy and the protection of their personal data. The second is that  
3 Defendants’ use of licensed code as training data for Codex and Copilot harms Plaintiffs’ property  
4 interests in the licensed code because Copilot already has or likely will reproduce Plaintiffs’ code  
5 in violation of those licenses.

### 6 **1. Injury to Privacy Rights**

7 Defendants argue that Plaintiffs fail to plausibly allege any disclosure of personal  
8 information, and therefore fail to allege an actual or imminent injury sufficient to confer standing.  
9 In the complaint, Plaintiffs do not identify the specific sensitive or private information at issue.  
10 Plaintiffs therefore do not allege facts sufficient for the Court to infer that Plaintiffs’ privacy  
11 interests are implicated by the alleged misuse of such information. *See Ji v. Naver Corp.*, No. 21-  
12 cv-05143-HSG, 2022 WL 4624898, at \*7 (N.D. Cal. Sept. 30, 2022) (dismissing privacy-based  
13 claim for lack of standing where plaintiffs “have not alleged enough facts to show that the  
14 [personal data at issue] are the type of information that could give rise to a privacy injury”); *I.C. v.*  
15 *Zynga, Inc.*, 600 F. Supp. 3d 1034, 1049 (N.D. Cal. Apr. 29, 2022) (evaluating whether privacy  
16 interests were implicated by allegedly disclosed personal information to determine whether  
17 plaintiffs had standing).

18 Plaintiffs have not met their burden to allege facts demonstrating an injury-in-fact  
19 sufficient to confer standing for their privacy-based claims. Plaintiffs’ claims for breach of the  
20 GitHub Privacy Policy and Terms of Service, violation of the CCPA, and negligence are  
21 dismissed with leave to amend.

### 22 **2. Injury to Property Rights**

23 Defendants argue that Plaintiffs do not identify any instance of Copilot reproducing  
24 Plaintiffs’ licensed code and therefore fail to plead a particularized injury sufficient to confer  
25 standing.

26 “For an injury to be ‘particularized,’ it ‘must affect the plaintiff in a personal and  
27 individual way.’” *Spokeo*, 578 U.S. at 339 (quoting *Lujan*, 504 U.S. at 560 n.1). “[A]t an  
28 irreducible minimum, Art. III requires the party who invokes the court’s authority to ‘show that he

1 personally has suffered some actual or threatened injury.” *Valley Forge Christian Coll. v. Ams.*  
2 *United for Separation of Church and State, Inc.*, 454 U.S. 464, 472 (1982) (quoting *Gladstone*  
3 *Realtors v. Village of Bellwood*, 441 U.S. 91, 99 (1979)).

4 Plaintiffs argue that they “have alleged the requisite particularized injury” because they  
5 “allege that Defendants have violated provisions of [the] open-source licenses” under which  
6 Plaintiffs published their code. ECF No. 67 at 14. In the complaint, Plaintiffs allege that  
7 Defendants violate those licenses because Copilot reproduces Plaintiffs’ code as output with  
8 missing or incorrect attribution, copyright notices, and license terms.<sup>7</sup> However, while Plaintiffs  
9 identify several instances in which Copilot’s output matched licensed code written by a Github  
10 user, Compl. ¶¶ 56, 71, 74, 87-89, none of these instances involve licensed code published to  
11 GitHub by Plaintiffs. Because Plaintiffs do not allege that they themselves have suffered the  
12 injury they describe, they do not have standing to seek retrospective relief for that injury.<sup>8</sup>

13 Plaintiffs also argue that they “also allege an imminent and significant harm—now that  
14 their software code is part of the training data . . . , [it] can be output at any time without the  
15 legally necessary notices and attribution, leading to increased and imminent risk of  
16 misappropriation.” ECF No. 67 at 15. To the extent that Plaintiffs allege an injury based on the  
17 increased risk of misappropriation of their own licensed code, that injury is sufficiently  
18 particularized. However, an increased risk of future harm alone is not sufficiently concrete to  
19 confer standing for damages. *TransUnion*, 141 S. Ct. at 2210-11 (finding “persuasive” argument

---

21 <sup>7</sup> In opposition, Plaintiffs suggest they were also injured by Defendants’ use of Plaintiffs’ licensed  
22 code as training data for Codex and Copilot. ECF No. 66 at 12, 21. But Plaintiffs’ complaint does  
23 not describe such an injury. In the complaint, Plaintiffs allege that “Defendants had access to but  
24 were not licensed by Plaintiffs . . . to train any . . . functional prediction engine using the Licensed  
25 Materials.” Compl. ¶ 149; *see also id.* ¶ 37 (describing one question common to the putative class  
26 as “[w]hether Defendants violated the Licenses governing use of the Licensed Materials by using  
27 them to train Copilot”). However, Plaintiffs do not actually allege that Defendants’ use of  
28 licensed code to train Codex and Copilot constituted a breach of the open-source licenses at issue.  
*See id.* ¶¶ 172-87 (alleging that Defendants breached the licenses by failing to provide attribution  
in output, failing to provide copyright notices in output, and failing to identify applicable licenses  
and the text of those licenses in output). Because Plaintiffs do not allege they were injured by  
Defendants’ use of licensed code as training data, the Court does not address whether such an  
injury is sufficient to confer standing.

<sup>8</sup> This problem is related to redressability: Plaintiffs allege no injury redressable by the monetary  
damages they seek.



1 that, “in a suit for damages, the mere risk of future harm . . . cannot qualify as a concrete harm . . .  
 2 unless the exposure to the risk of future harm itself causes a *separate* concrete harm”) (emphasis  
 3 in original); *id.* at 2213 (holding that “the risk of future harm on its own does not support Article  
 4 III standing for the plaintiffs’ damages claims”). Plaintiffs do not allege any additional, concrete  
 5 harm associated with this increased risk of misappropriation. Thus, an increased risk of  
 6 misappropriation cannot provide standing for Plaintiffs’ damages claims.

7 Finally, Plaintiffs argue that, “[g]iven the number of times users may use Copilot, it is a  
 8 virtual certainty [that] any particular plaintiff’s code will be displayed either with copyright  
 9 notices removed or in violation of Plaintiffs’ open-source licenses for profit.” ECF No. 67 at 15.  
 10 Though this does not support standing for retrospective damages – for which, as discussed above,  
 11 Plaintiffs must allege that they have suffered a qualifying injury-in-fact that has actually occurred  
 12 to them – it may support standing for injunctive relief.

13 “[A] person exposed to a risk of future harm may pursue forward-looking, injunctive relief  
 14 to prevent the harm from occurring, at least so long as the risk of harm is sufficiently imminent  
 15 and substantial.” *TransUnion*, 141 S. Ct. at 2210. “An allegation of future injury may suffice if  
 16 the threatened injury is ‘certainly impending,’ or there is a ‘substantial risk that the harm will  
 17 occur.’” *Susan B. Anthony List v. Driehaus*, 573 U.S. 149, 158 (2014) (quoting *Clapper v.*  
 18 *Amnesty Int’l USA*, 568 U.S. 414 n.5 (2013)). A substantial risk means there is a “realistic danger  
 19 of sustaining a direct injury.” *Pennell v. City of San Jose*, 485 U.S. 1, 8 (1988) (quoting *Babbitt v.*  
 20 *United Farm Workers Nat’l Union*, 442 U.S. 289, 298 (1979)).

21 Plaintiffs plausibly allege that there is at least a substantial risk that Defendants’ programs  
 22 will reproduce Plaintiffs’ licensed code as output. Plaintiffs allege that both Codex and Copilot  
 23 were trained on data that included all public GitHub repositories. Plaintiffs further allege that the  
 24 programs reproduce well-known code in response to related prompts and offer several concrete  
 25 examples of such reproduction. Finally, Plaintiffs allege that GitHub’s own internal research  
 26 shows that Copilot reproduces code from training data “about 1% of the time.” Compl. ¶¶ 56, 71,  
 27 74, 87-90. Plaintiffs thus plausibly allege that, absent injunctive relief, there is a realistic danger  
 28 that Codex or Copilot will reproduce Plaintiffs’ licensed code as output. Plaintiffs further allege

1 that Defendants have modified Copilot to ensure that it no longer reproduces license text,  
2 attribution, and copyright notices. Taking the facts of the complaint as true and construing all  
3 inferences in the Plaintiffs’ favor, the Court can reasonably infer that, should Plaintiffs’ code be  
4 reproduced as output, it will be reproduced in a manner that violates the open-source licenses  
5 under which Plaintiffs published their code.<sup>9</sup>

6 While Plaintiffs have failed to establish an injury-in-fact sufficient to confer standing for  
7 their claims for damages based on injury to property rights, they have standing to pursue  
8 injunctive relief on such claims.

9 **B. Proceeding Pseudonymously**

10 Defendants argue that the complaint should be dismissed because Plaintiffs are not entitled  
11 to proceed under fictitious names.<sup>10</sup>

12 “The normal presumption in litigation is that parties must use their real names.” *Doe v.*  
13 *Kamehameha Schs./Bernice Pauahi Bishop Est.*, 596 F.3d 1036, 1042 (9th Cir. 2010). Parties  
14 may proceed pseudonymously only “in the ‘unusual case’ when nondisclosure of the party’s  
15 identity ‘is necessary . . . to protect a person from harassment, injury, ridicule or personal  
16 embarrassment.” *Does I Thru XXIII v. Advanced Textile Corp.*, 214 F.3d 1058, 1067-68 (9th Cir.  
17 2000) (alteration in original) (quoting *United States v. Doe*, 655 F.2d 920, 922 n.1 (9th Cir.  
18 1981)).

19 Threats of physical harm “present[] the paradigmatic case for allowing anonymity.”  
20 *Kamehameha*, 596 F.3d at 1043. Where a party seeks to proceed pseudonymously on the basis of

21 \_\_\_\_\_  
22 <sup>9</sup> Defendants argue that Plaintiffs must also allege facts regarding the type of code they published,  
23 what problem that code solves, how frequently that code might appear on GitHub, and how likely  
24 it is that a user would enter a prompt that would generate a match to that code. ECF No. 50 at 16.  
25 Defendants suggest that, absent such facts, Plaintiffs’ alleged injury is too “conjectural or  
26 hypothetical” to confer standing. *Id.* (quoting *Lujan*, 504 U.S. at 550). Defendants ask too much  
27 of Plaintiffs at the pleading stage. Such facts might help Defendants—the only parties with  
28 knowledge of how Copilot and Codex were designed and operate—understand whether Plaintiffs’  
alleged injury has already occurred or how soon it is likely to occur. However, taking the facts in  
the complaint as true, and construing all inferences in Plaintiffs’ favor, their alleged future injury  
is neither conjectural nor hypothetical: their licensed code was used to train these programs and  
there is presently a realistic danger that the code will be reproduced as output.

<sup>10</sup> Because the caselaw uses “pseudonymous” and “anonymous” interchangeably, this Court does  
so as well.

1 retaliatory harm, “a district court must balance five factors: ‘(1) the severity of the threatened  
2 harm, (2) the reasonableness of the anonymous party’s fears, . . . (3) the anonymous party’s  
3 vulnerability to such retaliation,’ (4) the prejudice to the opposing party, and (5) the public  
4 interest.” *Id.* at 1042 (quoting *Advanced Textile*, 214 F.3d at 1068).

5 Plaintiffs seek to proceed pseudonymously because, “through their counsel, [they] have  
6 received legitimate and credible threats of physical violence.” ECF No. 67 at 17. Plaintiffs  
7 highlight three threatening emails sent to counsel prior to the filing of the operative complaint. All  
8 three emails contain veiled threats wishing death upon Plaintiffs’ counsel on the basis of his  
9 involvement in this lawsuit. ECF No. 68-1 at 2 (“[I]magine shooting against [AI] . . . . [G]o kys. .  
10 . . [K]ill urself.”); ECF No. 68-2 at 2 (“I hope you f\*cking die you piece of sh\*t. It’s people like  
11 you why this world sucks so f\*cking bad and we can’t have nice things. I literally hope someone  
12 murder [sic] you. Go f\*cking die.”); ECF No. 68-3 at 2 (“[G]o f\*cking cry about [G]it[H]ub you  
13 f\*cking piece of sh\*t n\*\*\*\*r, [I] hope your throat gets cut open and every single family member  
14 of you [sic] is burnt to death.”). Plaintiffs’ counsel have received “many” such messages. ECF  
15 No. 67 at 17. Plaintiffs argue they should be permitted to proceed pseudonymously because they  
16 reasonably fear they will be subject to threats of retaliation and violence if their identities are  
17 disclosed.

18 The first two factors of the balancing test for retaliatory harm – severity of the threatened  
19 harm and the reasonableness of the fear of such harm – “are intricately related and should be  
20 addressed together.” *Kamehameha*, 596 F.3d at 1040. Plaintiffs need not prove the speakers  
21 “intend to carry out the threatened retaliation,” but rather “that a reasonable person would believe  
22 that the threat might actually be carried out.” *Advanced Textile*, 214 F.3d at 1071. “It is in the  
23 particular purview of the district court to view alleged threats in context and determine what the  
24 ‘reasonable’ person in the plaintiffs’ situation would fear.” *Kamehameha*, 596 F.3d at 1044. “In  
25  
26  
27  
28

1 context, a plaintiff might reasonably fear a veiled threat of violence.” *Id.*<sup>11</sup>

2 The threatened harm in this case – death – is plainly severe. *Id.* at 1043 (describing threats  
3 of physical retaliation, including death, as “undoubtedly severe”); *Doe v. Steagall*, 653 F.2d 180,  
4 186 (5th Cir. 1981) (permitting plaintiffs to proceed pseudonymously where they faced “threats of  
5 violence,” including veiled death threats, “generated by this case”); *Doe v. Univ. Acct. Serv., LLC*,  
6 No. 09-CV-01563-BAS-JLB, 2022 WL 623913, at \*4 (S.D. Cal. Mar. 3, 2022) (permitting  
7 plaintiff to proceed pseudonymously where he received death threats). While the threatening  
8 emails were not sent to Plaintiffs directly, the emails wish death upon Plaintiffs’ counsel on the  
9 basis of their involvement in this lawsuit. It is reasonable for Plaintiffs to fear that such threats  
10 might be carried out against them if their identities were to become public.

11 GitHub and Microsoft suggest that Plaintiffs’ fear is unreasonable because “the types of  
12 nasty messages at issue here are a fact of modern life in the era of internet ‘trolls.’” ECF No. 72 at  
13 19. GitHub and Microsoft do not explain why the rise of internet trolls renders Plaintiffs’ fears of  
14 harm unreasonable. *Cf. United States v. Bagdasarian*, 652 F.3d 1113, 1126-27 (9th Cir. 2011)  
15 (Wardlaw, J., concurring in part) (describing the connection between anonymous internet posts  
16 and subsequent real-world violence). Sending direct messages containing veiled death threats  
17 would seem to constitute behavior beyond trolling. *See Merriam Webster’s Collegiate Dictionary*  
18 1341 (11th ed. 2003) (defining “troll,” in relevant part, as “to antagonize (others) online by  
19 *posting* inflammatory, irrelevant, or offensive comments or content” and “to harass, criticize, or  
20 antagonize (someone) esp[ecially] by provocatively disparaging or mocking *public* statements,  
21 postings, or acts”) (emphasis added). These were not public posts mocking or antagonizing  
22 Plaintiffs’ counsel, but rather private emails, sent directly to Plaintiffs’ counsel, wishing him and  
23 his family violent death. Where many individuals take the time to send private, threatening  
24

---

25 <sup>11</sup> Of note, the Ninth Circuit has explained that the standard for reasonable fear in this context is  
26 more permissive than in First Amendment “true threat” cases. *Kamehameha*, 596 F.3d at 1044 n.6  
27 (“The First Amendment cases discussing the concept of a ‘true threat’ . . . pose a higher bar to  
28 finding a reasonable fear. In those cases, one party’s fear of the threat must be weighed against  
the opposing party’s first amendment right to speak freely because the threatened party seeks to  
prevent the other party’s speech.”).

1 emails, it is reasonable to fear that some of those individuals might carry out their threats. The  
2 Court finds that the first two factors weigh in favor of permitting Plaintiffs to proceed  
3 pseudonymously.

4 Defendants identify no prejudice from Plaintiffs proceeding pseudonymously. The Ninth  
5 Circuit has explained that, where a defendant does not know the plaintiff's name, "at some later  
6 point in the proceedings it may be necessary to reveal plaintiffs' identities to defendants so that  
7 defendants may refute [their] individualized accusations." *Advanced Textile*, 214 F.3d at 1058.  
8 "But where the defendants know the plaintiffs' names, 'anonymity need not, and should not,  
9 impede either party's ability to develop its case' even though it is 'foreseeable that anonymity  
10 [will] raise problems for discovery.'" *Doe 1 v. Nat'l Collegiate Athletic Assoc.*, No. 22-cv-01559-  
11 LB, 2022 WL 3974098, at \*2 (N.D. Cal. Aug. 30, 2022) (quoting *Jane Roes 1-2 v. SFBSC Mgmt.,*  
12 *LLC*, 77 F. Supp. 3d 990, 996 (N.D. Cal. 2015)). Plaintiffs have disclosed their true names to  
13 Defendants subject to a protective order, so pseudonymity should not impede Defendants' ability  
14 to develop their case. Though pseudonymity may pose certain logistical challenges during  
15 discovery, this case remains at the pleadings stage. *See Doe v. County of El Dorado*, No. 2:13-  
16 CV-01433-KJM, 2013 WL 6230342, at \*5 (E.D. Cal. Dec. 2, 2013) (explaining that, "at the  
17 pre-discovery stage . . . the court need not yet consider the prejudice defendant will suffer during  
18 discovery," as "the relevant prejudice is that which defendant presently suffers as a result of  
19 plaintiff's anonymity"). Defendants do not articulate any prejudice associated with Plaintiffs  
20 proceeding pseudonymously at this stage of the litigation, and this Court is not aware of any.  
21 Therefore, the Court finds that the prejudice factor does not weigh against permitting Plaintiffs to  
22 proceed pseudonymously at this stage.

23 Finally, the public interest factor does not weigh against permitting Plaintiffs to proceed  
24 pseudonymously. "The normal presumption in litigation . . . that parties must use their real names  
25 . . . is loosely related to the public's right to open courts and the right of private individuals to  
26 confront their accusers." *Kamehameha*, 596 F.3d at 1042. Where the plaintiffs' identities are not  
27 central to the issues raised by a case, however, the public interest may not be harmed by  
28 permitting plaintiffs to proceed pseudonymously. *See Advanced Textile*, 214 F.3d at 1072

1 (reversing denial of anonymity where “[t]he district court did not explain, and we fail to see, how  
2 disguising plaintiffs’ identities will obstruct public scrutiny of the important issues in th[e] case”);  
3 *Kamehameha*, 596 F.3d at 1043 (noting that, where plaintiffs brought “claims of widespread  
4 discrimination,” “it [wa]s difficult to see ‘how disguising plaintiffs’ identities w[ould] obstruct  
5 public scrutiny of the important issues in th[e] case’”) (quoting *Advanced Textile*, 214 F.3d at  
6 1072). Withholding the true identities of the individual software developers who bring this case  
7 will not obstruct public scrutiny of the issues raised. Plaintiffs’ names and identities have no  
8 bearing on the central issues of this case, including whether Defendants’ AI-based coding tools  
9 illegally reproduce licensed code used as training data. The Court finds that the public interest  
10 factor does not weigh against anonymity at this stage of the litigation. *See El Dorado*, 2013 WL  
11 6230342, at \*6 (finding public interest factor weighed in favor of anonymity where the  
12 “[p]laintiff’s identity appears to have no bearing on the resolution of the issues, and a pseudonym  
13 will not impede public access to the substance of the proceedings”).<sup>12</sup>

14 Balancing these factors, the Court finds that Plaintiffs may proceed pseudonymously at this  
15 time.

### 16 C. Rule 8(a)

17 Defendants argue that the complaint should be dismissed under Federal Rule of Civil  
18 Procedure 8(a) because Plaintiffs do not plead sufficient facts regarding the role of each Defendant  
19 in the alleged misconduct.

20 A complaint must contain “a short and plain statement of the claim showing that the  
21 pleader is entitled to relief.” Fed. R. Civ. P. 8(a)(2). “[T]he ‘short and plain statement’ must  
22 provide the defendant with ‘fair notice of what the plaintiff’s claim is and the grounds upon which  
23 it rests.’” *Dura Pharms., Inc. v. Broudo*, 544 U.S. 336, 346 (2005) (quoting *Conley v. Gibson*, 355  
24 U.S. 41, 47 (1957)). To comply with Rule 8(a), a plaintiff “must allege the basis of his claim  
25

---

26 <sup>12</sup> The Court is mindful that, like the other factors, the public interest in Plaintiffs’ identities may  
27 change as the suit progresses. *See Doe v. NFL Enters., LLC*, No. C 17-00496 WHA, 2017 WL  
28 697420, at \*2 (N.D. Cal. Feb. 22, 2017) (noting that “class members will . . . have a right to know  
the identity of their representative in this litigation” in later stages of the case).

1 against each defendant . . . to put defendants on sufficient notice of the allegations against them.”  
 2 *Gauvin v. Trombatore*, 682 F. Supp. 1067, 1071 (N.D. Cal. 1988).

3 The OpenAI Defendants argue that Plaintiffs plead claims against all six OpenAI entities  
 4 without specifying the role or conduct of each entity, thus violating Rule 8(a). But Plaintiffs  
 5 allege that two of the OpenAI Defendants – OpenAI, Inc. and OpenAI, L.P. – “programmed,  
 6 trained, and maintain[]” Codex, and that OpenAI, Inc. owns and controls all of the other OpenAI  
 7 Defendants. Compl. ¶ 23. Plaintiffs plead additional specific facts regarding the ownership and  
 8 control of each of the OpenAI Defendants. *See, e.g., id.* ¶ 24 (“OpenAI, L.P. is a wholly owned  
 9 subsidiary of OpenAI, Inc. . . . [and] the OpenAI entity that co-created Copilot and offers it jointly  
 10 with GitHub.”); *id.* ¶ 25 (“OpenAI GP is the general partner of OpenAI, L.P. OpenAI GP  
 11 manages and operates the day-to-day business and affairs of OpenAI, L.P. . . . OpenAI GP was  
 12 aware of the unlawful conduct alleged herein and exercised control over OpenAI, L.P.”). Such  
 13 allegations sufficiently put each OpenAI entity on notice of the basis on which Plaintiffs allege it  
 14 may be liable for the challenged conduct.<sup>13</sup>

15 Similarly, Microsoft and GitHub argue that Plaintiffs’ complaint fails to satisfy Rule 8(a)  
 16 because Plaintiffs do not allege which specific acts taken by each Defendant violate the DMCA.  
 17 But Plaintiffs allege facts which sufficiently provide notice of the basis on which each GitHub and  
 18 Microsoft are sued. Microsoft is alleged to maintain an ownership interest in GitHub and OpenAI,  
 19 L.P., which allegedly co-created Copilot. Compl. ¶ 22 (“Microsoft owns and operates GitHub.  
 20 Through its corporate ownership, control of the GitHub Board of Directors, active management,  
 21

---

22  
 23 <sup>13</sup> The Court does not decide here whether such alleged relationships in fact provide a sufficient  
 24 basis for liability against each of the OpenAI entities for any of the pleaded claims, because that  
 25 issue is not presently before the Court. The function of Rule 8 in this context is to “give  
 26 Defendants fair notice of the allegations against them,” *Ketayi v. Health Enrollment Grp.*, 516 F.  
 27 Supp. 3d 1092, 1120 (S.D. Cal. 2021) (citations omitted), whereas the purpose of a Rule 12(b)(6)  
 28 motion is to “test[] the legal sufficiency of the pleadings and allow[] a court to dismiss a complaint  
 upon a finding that the plaintiff has failed to state a claim upon which relief may be granted,”  
*Lessin v. Ford Motor Co.*, 600 F. Supp. 3d 1137, 1141 (S.D. Cal. 2022) (citation omitted). *See*  
*also Olson v. Puckett*, No. 221CV01482KJMDMC, 2023 WL 2602174, at \*1 (E.D. Cal. Mar. 22,  
 2023) (“To the extent there is a difference between the standards imposed by Rule 8 and Rule  
 12(b)(6), this court interprets the Rule 8 requirement as the lesser of the two.”).

1 and other means, Microsoft sells, markets and distributes Copilot.”); *id.* ¶ 7 (“Microsoft obtained a  
 2 partial ownership interest in OpenAI . . . . As OpenAI’s largest investor . . . Microsoft exerts  
 3 considerable control over OpenAI.”); *id.* ¶ 24 (“OpenAI, L.P. is the OpenAI entity that co-created  
 4 Copilot and offers it jointly with GitHub.”). Plaintiffs further allege, on information and belief,  
 5 that “Microsoft, GitHub, [and] OpenAI . . . have worked together to create Copilot.” *Id.* ¶ 241.  
 6 Such allegations provide sufficient notice of the basis of Plaintiffs’ DMCA claims against each  
 7 Defendant.<sup>14</sup> The complaint is not deficient under Rule 8(a).

8 **D. Rule 12(b)(6)**

9 Defendants move to dismiss most of Plaintiffs’ claims for failure to state a claim on which  
 10 relief may be granted. In opposition, Plaintiffs do not address Defendants’ substantive arguments  
 11 about the sufficiency of Plaintiffs’ claims for violation of Section 1202(a) of the DMCA, tortious  
 12 interference, fraud, false designation of origin, or violation of the CCPA. Plaintiffs instead state  
 13 that they “do not concede” these claims “have been inadequately pled,” but request leave to amend  
 14 the claims. ECF No. 66 at 30; ECF No. 67 at 32 n.17. Without briefing from both parties, the  
 15 Court will not evaluate the merits of Defendants’ arguments as to those claims.

16 Mindful that leave to amend should be freely granted, the Court dismisses Plaintiffs’  
 17 claims for violation of Section 1202(a) of the DMCA, tortious interference, fraud, false  
 18 designation of origin, and violation of the CCPA with leave to amend.

19 **1. Copyright Preemption**

20 Defendants argue that several of Plaintiffs’ state law claims are preempted by Section 301  
 21 of the Copyright Act. Because most of these claims were previously dismissed, the Court only  
 22 considers whether Plaintiffs’ unjust enrichment claim is subject to copyright preemption.

23 The Copyright Act of 1976 expressly preempts state claims where the plaintiff’s work  
 24 “come[s] within the subject matter of copyright” and the state law grants “legal or equitable rights  
 25 that are equivalent to any of the exclusive rights within the general scope of copyright.” 17 U.S.C.  
 26 § 301(a). The Ninth Circuit has established a two-part test to determine whether state law claims

27 \_\_\_\_\_  
 28 <sup>14</sup> Again, the Court does not determine whether such allegations are sufficient to plead a basis for liability against any Defendant. *See supra* note 13.



1 are preempted. First, “[the court must] decide ‘whether the “subject matter” of the state law claim  
2 falls within the subject matter of copyright as described in 17 U.S.C. §§ 102 and 103.’” *Maloney*  
3 *v. T3Media, Inc.*, 853 F.3d 1004, 1010 (9th Cir. 2017). If it does, the court must “determine  
4 ‘whether the rights asserted under state law are equivalent to the rights contained in 17 U.S.C.  
5 § 106, which articulates the exclusive rights of copyright holders.” *Id.* “If a state law claim  
6 includes an ‘extra element’ that makes the right asserted qualitatively different from those  
7 protected under the Copyright Act, the state law is not preempted by the Copyright Act.” *Altera*  
8 *Corp. v. Clear Logic, Inc.*, 424 F.3d 1079, 1089 (9th Cir. 2005).

9 Plaintiffs’ opposition generally does not address Defendants’ preemption arguments.<sup>15</sup>  
10 However, Plaintiffs do argue generally that their state law claims are qualitatively different from  
11 claims under the Copyright Act because they are not solely about the unauthorized reproduction of  
12 Plaintiffs’ code, but also the unauthorized use of such code. Plaintiffs suggest that state law tort  
13 claims concerning unauthorized use are not preempted by the Copyright Act, and that “Plaintiffs  
14 allege that Defendants, through their unauthorized *use* of Plaintiffs’ code to train Codex and  
15 Copilot, and their display of Plaintiffs’ code to others for commercial gain, violated Plaintiffs’  
16 rights under state and common law.” ECF No. 66 at 18 (emphasis added).

17 Plaintiffs are correct that state law tort claims concerning unauthorized use are not  
18 preempted by the Copyright Act. “The [exclusive] rights protected under the Copyright Act  
19 include the rights of reproduction, preparation of derivative works, distribution, and display.”  
20 *Ryan v. Editions Ltd. W., Inc.*, 786 F.3d 754, 760 (9th Cir. 2015) (alteration in original) (quoting  
21 *Altera*, 424 F.3d at 1089). “A state law tort claim concerning the unauthorized use of the  
22 software’s end-product is not within the rights protected by the federal Copyright Act.” *Altera*,  
23 424 F.3d at 1079. However, in their complaint, Plaintiffs do not allege that Defendants’ use of  
24 Plaintiffs’ code for training purposes violated their rights. Rather, Plaintiffs base their unjust

---

25  
26 <sup>15</sup> Plaintiffs focus this part of their opposition on their breach of open-source license claims, which  
27 Defendants do not move to dismiss on copyright preemption grounds. ECF No. 66 at 17-18.  
28 Plaintiffs additionally argue that their negligence claim and UCL unlawful conduct claim based on  
violation of the DMCA are not preempted. *Id.* Defendants did not move to dismiss Plaintiffs’  
negligence claims on copyright preemption grounds, and only moved to dismiss Plaintiffs’ UCL  
claims to the extent they were based on preempted state law claims.

1 enrichment claim on Defendants’ reproduction of Plaintiffs’ code as output and Defendants’  
 2 preparation of derivative works, both of which are rights protected under the Copyright Act.  
 3 Compl. ¶¶ 204-09 (pleading unjust enrichment claim based on profit derived from both  
 4 reproduction and preparation of derivative works). Because the rights on which Plaintiffs base  
 5 their unjust enrichment claim are within those protected by the federal Copyright Act, Plaintiffs’  
 6 unjust enrichment claim is subject to preemption.

7 Plaintiffs’ claim for unjust enrichment is dismissed with leave to amend.

## 8 2. DMCA Section 1202(b) Claim

9 Because the Court previously dismissed Plaintiffs’ claim under Section 1202(a) of the  
 10 DMCA, the Court now only considers the sufficiency of Plaintiffs’ Section 1202(b) claim.

11 “Copyright law restricts the removal or alteration of copyright management information  
 12 (“CMI”) – information such as the title, the author, the copyright owner, the terms and conditions  
 13 for use of the work, and other identifying information set forth in a copyright notice or conveyed  
 14 in connection with the work.” *Stevens v. Corelogic, Inc.*, 899 F.3d 666, 671 (9th Cir. 2018).  
 15 Section 1202(b) of the DMCA provides that one cannot, without authority, (1) “intentionally  
 16 remove or alter any” CMI, (2) “distribute . . . [CMI] knowing that the [CMI] has been removed or  
 17 altered,” or (3) “distribute . . . copies of works . . . knowing that [CMI] has been removed or  
 18 altered” while “knowing, or . . . having reasonable grounds to know, that it will induce, enable,  
 19 facilitate, or conceal” infringement. 17 U.S.C. § 1202(b).

20 Plaintiffs allege that their licensed code contains CMI including copyright notices, titles,  
 21 authors’ names, copyright owners’ names, terms and conditions for use of the code, and  
 22 identifying numbers or symbols. Compl. ¶ 144. Plaintiffs allege that Defendants removed or  
 23 altered that CMI from licensed code, distributed CMI knowing CMI had been removed or altered,  
 24 and distributed copies of the code knowing that CMI had been removed or altered, all while  
 25 knowing and possessing reasonable grounds to know that doing so would induce infringement. *Id.*  
 26 ¶¶ 148, 153-55, 157.

27 Defendants argue that Plaintiffs have not plausibly alleged claims under Sections  
 28 1202(b)(1) or (b)(3) because these provisions require some active conduct that removes or alters

1 CMI. Defendants argue that the complaint merely alleges “the passive non-inclusion of CMI” by  
 2 neutral technology which excerpts code without the accompanying CMI, rather than the active  
 3 removal of CMI from licensed code. ECF No. 50 at 22. This semantic distinction is not  
 4 meaningful. Plaintiffs allege that the relevant CMI was affixed to their licensed code and that  
 5 Defendants were aware that such CMI appeared repeatedly across the data used to train Codex and  
 6 Copilot. Compl. ¶ 92 (CMI “usually appears just before a given block of code”); *id.* ¶ 94 (“[I]n a  
 7 blog post, GitHub noted[,] ‘In one instance, GitHub Copilot suggested starting an empty file with  
 8 something it had even seen more than a whopping 700,000 different times during training—that  
 9 was the GNU General Public License.’”). Defendants subsequently trained these programs to  
 10 ignore or remove CMI and therefore stop reproducing it. *Id.* ¶¶ 94-95 (“As GitHub explains:  
 11 ‘GitHub Copilot *has* changed to require a minimum file content.’ . . . Copilot no longer reproduces  
 12 these types of CMI . . . . It has been altered not to.”) (emphasis in original). Defendants knew that  
 13 these programs reproduced training data as output. *Id.* ¶ 90. Plaintiffs thus plead sufficient facts  
 14 to support a reasonable inference that Defendants intentionally designed the programs to remove  
 15 CMI from any licensed code they reproduce as output.<sup>16</sup>

16 Defendants further argue that Plaintiffs do not sufficiently plead scienter. As the Ninth  
 17 Circuit has explained, “the mental state requirement in Section 1202(b) must have a more specific  
 18 application than the universal possibility of encouraging infringement.” *Stevens*, 899 F.3d at 674.  
 19 *Stevens* involved software that allegedly removed metadata from image files which the defendant  
 20 subsequently distributed in violation of the DMCA. The Ninth Circuit affirmed the grant of  
 21 summary judgment to the defendant because the plaintiffs had not offered any evidence that the  
 22 removal of CMI would impair their policing of infringement. *Id.* at 675. At summary judgment,  
 23 “specific allegations as to how identifiable infringements ‘will’ be affected are necessary”; a  
 24 plaintiff “must make an affirmative showing, such as by demonstrating a past ‘pattern of conduct’  
 25 or ‘modus operandi,’ that the defendant was aware or had reasonable grounds to be aware of the  
 26 probable future impact of its actions.” *Id.* at 674.

27 \_\_\_\_\_  
 28 <sup>16</sup> In other words, Plaintiffs allege that Defendants designed “neutral” programs to effectuate the  
 “non-inclusion” of CMI by actively removing it from Plaintiffs’ licensed code.

1 On a motion to dismiss, a plaintiff must allege sufficient facts to support the reasonable  
2 inference that the defendant “knew or had a reasonable basis to know that the removal or alteration  
3 of CMI . . . w[ould] aid infringement.” *Harrison v. Pinterest, Inc.*, No. 20-cv-05290-EJD, 2022  
4 WL 4348460, at \* 5 (N.D. Cal. Sept. 19, 2022). At the pleading stage, mental conditions  
5 generally need not be alleged with specificity. “Federal Rule of Civil Procedure 9(b) provides that  
6 ‘intent, knowledge, and other conditions of a person’s mind may be alleged generally.’ Language  
7 in *Stevens* . . . does not indicate otherwise; there, the Ninth Circuit held that ‘a plaintiff bringing a  
8 Section 1202(b) claim must make an affirmative showing’ of scienter in the *summary judgment*  
9 context.” *Logan v. Meta Platforms, Inc.*, --- F. Supp. 3d ---, 2022 WL 14813836, at \*9 (N.D. Cal.  
10 Oct. 25, 2022) (emphasis in original) (internal citation omitted) (quoting *Stevens*, 899 F.3d at  
11 674); *see also Izmo, Inc. v. Roadster, Inc.*, No. 18-cv-06092-NC, 2019 WL 13210561, at \*4 (N.D.  
12 Cal. Mar. 26, 2019) (“Whether [the defendant] knew or should have known that its activities  
13 would induce or enable an infringement of [the plaintiff’s] rights is more suited to summary  
14 judgment.”).

15 Plaintiffs allege that Defendants knew the code they used as training data for Codex and  
16 Copilot routinely contained CMI. Compl. ¶ 94. Plaintiffs also allege that GitHub knew that CMI  
17 was important for protecting copyright interests. GitHub regularly processed DMCA takedowns,  
18 such that it was aware its platform was used to distribute code with removed or altered CMI in a  
19 manner which induced infringement. *Id.* ¶ 165-67. Plaintiffs’ allegations raise the reasonable  
20 inference that Defendants knew or had reasonable grounds to know that removal of CMI carried a  
21 substantial risk of inducing infringement. *See, e.g., Schneider v. Youtube, LLC*, No. 20-cv-04423-  
22 JD, 2022 WL 3031212, at \*2 (N.D. Cal. Aug. 1, 2022) (finding plaintiffs sufficiently pleaded  
23 scienter where they alleged that the defendant knew files “routinely contain CMI, that CMI is  
24 valuable for protecting copyright holders, and that the distribution of works with missing CMI on  
25 [defendant’s platform] has induced . . . infringement,” supporting a “plausible inference” that  
26 defendant removed CMI “with knowledge that doing so carried a ‘substantial risk’ of inducing  
27 infringement”) (quoting *Stevens*, 899 F.3d at 676); *Batra v. PopSugar*, No. 18-cv-03752-HSG,  
28 2019 WL 482492, at \*2 (N.D. Cal. Feb. 7, 2019) (finding plaintiffs’ allegations regarding removal

1 of CMI supported a “plausible inference” that the defendant “kn[ew] that removing the CMI  
2 would help to conceal the alleged infringement”).

3 Defendants further argue that Plaintiffs fail to plausibly allege that Defendants distribute  
4 CMI “knowing the [CMI] has been removed or altered” in violation of Section 1202(b)(2).  
5 Plaintiffs’ opposition does not directly address this argument. In the complaint, Plaintiffs allege  
6 that “Defendants have a business practice of asserting and/or implying that Copilot is the author of  
7 the Licensed Materials” and that “Defendants’ false description of the source of Copilot’s Output  
8 facilitated or concealed infringement by Defendants and Copilot users.” Compl. ¶¶ 158-59.  
9 Plaintiffs do not identify the assertions, implications, and/or false descriptions of authorship or  
10 source at issue, nor do they plead facts that suggest such unidentified statements could constitute  
11 CMI. Plaintiffs separately allege that Copilot previously “would sometimes produce [CMI] . . .  
12 [which] was not always accurate,” but that Copilot “no longer reproduces these types of CMI,  
13 incorrect or otherwise, on a regular basis.” *Id.* ¶ 95. Plaintiffs plead no specific facts regarding  
14 the allegedly inaccurate CMI Copilot once produced alongside output, nor do they plead facts  
15 suggesting such inaccurate CMI is likely to be produced alongside their output in the future.  
16 Because the allegations in the complaint do not sufficiently allege the distribution of altered CMI,  
17 the Court finds that Plaintiffs have failed to state a claim under Section 1202(b)(2).

18 The Court denies Defendants’ motions to dismiss Plaintiffs’ claim under Sections  
19 1202(b)(1) and 1202(b)(3). Plaintiffs’ claim under Section 1202(b)(2) is dismissed with leave to  
20 amend.

### 21 3. Breach of License Claim

22 Defendants argue that Plaintiffs fail to plead the existence of a contract because they do not  
23 indicate which licenses are at issue or which provisions Defendants allegedly breached.<sup>17</sup>

24 Under California law, breach of contract requires plaintiffs to “plead ‘the contract,  
25 plaintiff’s performance (or excuse for nonperformance), defendant’s breach, and damage to  
26

---

27 <sup>17</sup> Defendants additionally argue that, to the extent Plaintiffs’ breach of license claim is based on  
28 Defendants’ training of Codex and Copilot using licensed code, it is foreclosed by GitHub’s  
Terms of Service. Because Plaintiffs’ breach of license claim is not based on such training, *see*  
Compl. ¶¶ 180-83, this argument is irrelevant.

1 plaintiff therefrom.” *Low v. LinkedIn Corp.*, 900 F. Supp. 2d 1010, 1028 (N.D. Cal. 2012)  
2 (quoting *Gautier v. Gen. Tel. Co.*, 234 Cal. App. 2d 302, 305 (1965)). “Identifying the specific  
3 provision of the contract allegedly breached by the defendant does not require the plaintiff to  
4 attach the contract or recite the contract’s terms verbatim. Rather, the plaintiff must identify with  
5 specificity the contractual obligations allegedly breached by the defendant.” *Williams v. Apple,*  
6 *Inc.*, 449 F. Supp. 3d 892, 908 (N.D. Cal. 2020) (quoting *Kaar v. Wells Fargo Bank, N.A.*, No. C  
7 16-01290 WHA, 2016 WL 3068396, at \*1 (N.D. Cal. June 1, 2016)).

8 Plaintiffs advance claims for breach of the eleven suggested licenses GitHub presents to  
9 users that require (1) attribution to the owner, (2) inclusion of a copyright notice, and (3) inclusion  
10 of the license terms. Compl. ¶ 34 n.4. Plaintiffs attach each of these licenses to the complaint.  
11 Plaintiffs allege that use of licensed code “is allowed only pursuant to the terms of the applicable  
12 Suggested License,” and that each such license requires that any derivative work or copy include  
13 attribution, a copyright notice, and the license terms. *Id.* ¶¶ 173, 34 n.4. Plaintiffs further allege  
14 that Codex and Copilot reproduce licensed code as output without attribution, copyright notice, or  
15 license terms, thereby violating the relevant provisions of each license. While Plaintiffs do not  
16 identify the specific subsections of each suggested license that correspond to each of these  
17 requirements, the Court finds that Plaintiffs have sufficiently identified “the contractual  
18 obligations allegedly breached,” as required to plead a breach of contract claim. *Williams*, 449 F.  
19 Supp. 3d at 908.

20 Defendants’ motions to dismiss Plaintiffs’ claim for breach of license is denied.

#### 21 4. Unfair Competition

22 Plaintiffs assert claims for unfair competition under the Lanham Act, the UCL, and  
23 California common law against GitHub and OpenAI Defendants. These claims are predicated on  
24 Plaintiffs’ claims for violation of the DMCA, tortious inference, false designation of origin,  
25 violation of the CCPA, and negligence.

26 Defendants move to dismiss Plaintiffs’ UCL claims for failure to sufficiently allege  
27 predicate claims. To the extent the predicate claims have been dismissed, Plaintiffs’ derivative  
28 UCL claims must also be dismissed with leave to amend. *See Eidmann v. Walgreen Co.*, 522 F.

1 Supp. 3d 634, 647 (N.D. Cal. 2021) (“If the ‘plaintiff cannot state a claim under the predicate law .  
2 . . . [the UCL] claim also fails.’”) (alterations in original) (quoting *Hadley v. Kellogg Sales Co.*, 243  
3 F. Supp. 3d 1074, 1094 (N.D. Cal. 2017)). Plaintiffs’ UCL claims predicated upon violation of  
4 Sections 1202(a) and 1202(b)(2) of the DMCA, tortious interference, false designation of origin,  
5 violation of the CCPA, and negligence are dismissed with leave to amend.

6 Plaintiffs’ UCL claims predicated on violation of Sections 1202(b)(1) and (b)(3) of the  
7 DMCA remain. Defendants argue that Plaintiffs fail to plead any economic injury arising from  
8 the predicate violation, as required for statutory standing under the UCL. To plead a cause of  
9 action under the UCL, a plaintiff must allege that the challenged conduct caused “some form of  
10 economic injury,” like “lost money or property.” *Kwikset Corp. v. Super. Ct.*, 51 Cal. 4th 310,  
11 323 (2011). As the California Supreme Court explained, a plaintiff may suffer economic injury by  
12 “hav[ing] a present or future property interest diminished” or “be[ing] deprived of money or  
13 property to which he or she has a cognizable claim.” *Id.*

14 Plaintiffs allege that they “have suffered monetary damages” as a result of all of the  
15 predicate violations listed in their UCL claim. Compl. ¶ 213. From the allegations in the  
16 complaint, however, the Court cannot discern how Defendants’ alleged violations of the DMCA  
17 have caused or will cause Plaintiffs economic injury. Plaintiffs’ opposition argues they “lost the  
18 value of [their] work, including their ability to receive compensation as well as the likelihood they  
19 would be retained or hired in the future,” and have suffered injury to their intellectual property  
20 rights, including “loss of value of the computer code and the value to code authors of their  
21 attribution rights.” ECF No. 66 at 25. However, no such injury is alleged in the complaint.  
22 Accordingly, Plaintiffs’ UCL claim predicated on violation of the DMCA is dismissed with leave  
23 to amend.

24 The OpenAI Defendants also move to dismiss Plaintiffs’ claims for unfair competition  
25 under the Lanham Act and California common law. Plaintiffs do not address these arguments in  
26 opposition, and therefore abandon the claims. *See Diamond S.J. Enter., Inc. v. City of San Jose*,  
27 No. 18-cv-01353-LHK, 2018 WL 5619746, at \*4 (N.D. Cal. Oct. 29, 2018) (“An opposition  
28 brief’s failure to address a motion to dismiss’[s] challenges to a claim constitutes abandonment of

1 that claim.”). Plaintiffs’ claims for unfair competition under the Lanham Act and California  
2 common law are dismissed with leave to amend.

### 3 5. Civil Conspiracy

4 Defendants argue that Plaintiffs’ civil conspiracy claim must be dismissed because civil  
5 conspiracy is not a standalone cause of action.

6 “Conspiracy is not a cause of action, but a legal doctrine that imposes liability on persons  
7 who, although not actually committing a tort themselves, share with the immediate tortfeasors a  
8 common plan or design in its perpetration.” *Applied Equip. Corp. v. Litton Saudi Arabia Ltd.*, 7  
9 Cal. 4th 503, 510-11 (1994) (in bank); *see also AccuImage Diagnostics Corp v. Terarecon, Inc.*,  
10 260 F. Supp. 2d 941, 947-48 (N.D. Cal. 2003) (dismissing standalone civil conspiracy claim with  
11 prejudice because “civil conspiracy is not a separate and distinct cause of action under California  
12 law” and explaining that any “amended [civil conspiracy] allegations . . . must be made within the  
13 sections of the complaint that contain plaintiff’s claims for the underlying” violations).

14 Because Plaintiffs cannot plead civil conspiracy as an independent cause of action, this  
15 claim must be dismissed with prejudice.

### 16 6. Declaratory Relief

17 Defendants move to dismiss Plaintiffs’ claim for declaratory relief, arguing that declaratory  
18 relief is not a standalone cause of action. Plaintiffs do not address Defendants’ arguments as to  
19 this claim in their opposition briefs, and therefore have abandoned the claim. *See Diamond S.J.*,  
20 2018 WL 5619746, at \*4.

21 “[D]eclaratory relief is not a standalone claim.” *Mayen v. Bank of Am. N.A.*, No. 14-cv-  
22 03757-JST, 2015 WL 179541, at \*5 (N.D. Cal. Jan. 14, 2015); *see also Sowinski v. Wells Fargo*  
23 *Bank, N.A.*, No. 11-6431-SC, 2012 WL 5904711, at \*1 (N.D. Cal. Nov. 26, 2012) (dismissing  
24 declaratory relief claim with prejudice and noting that “Plaintiff may still seek declaratory . . .  
25 relief in any further pleading, provided that he asserts a claim that could give rise to such relief”).  
26 Because declaratory relief is not a claim, granting leave to amend would be futile. Accordingly,  
27 Plaintiffs’ claim for declaratory relief is dismissed with prejudice.

28



**CONCLUSION**

Defendants’ motions to dismiss are granted in part and denied in part. Plaintiffs’ claims for violation of Sections 1202(a) and 1202(b)(2) of the DMCA, tortious interference in a contractual relationship, fraud, false designation of origin, unjust enrichment, unfair competition, breach of the GitHub Privacy Policy and Terms of Service, violation of the CCPA, and negligence are dismissed with leave to amend. Plaintiffs’ claims for civil conspiracy and declaratory relief are dismissed with prejudice.

Plaintiffs shall file an amended complaint within 28 days of this order.

**IT IS SO ORDERED.**

Dated: May 11, 2023

  
\_\_\_\_\_  
JON S. TIGAR  
United States District Judge

United States District Court  
Northern District of California

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28