1 ANNETTE L. HURST (SBN 148738)
   ahurst@orrick.com
2 ORRICK, HERRINGTON & SUTCLIFFE LLP
   405 Howard Street
3 San Francisco, CA 94105-2669
   Telephone: +1 415 773 5700
4 Facsimile: +1 415 773 5759

5 WILLIAM W. OXLEY (SBN 136793)
   woxley@orrick.com
6 ALYSSA CARIDIS (SBN 260103)
   acaridis@orrick.com
7 ORRICK, HERRINGTON & SUTCLIFFE LLP
   355 S. Grand Avenue
8 Los Angeles, CA 90071
   Telephone: +1 213 629 2020
9 Facsimile: +1 213 612 2499

10 *Attorneys for GitHub, Inc. and Microsoft Corporation*

11   UNITED STATES DISTRICT COURT

12   NORTHERN DISTRICT OF CALIFORNIA

13   OAKLAND DIVISION

14
   J. DOE 1, et al.,   Case No. 4:22-cv-6823-JST
15
             Individual and   Consolidated with Case No. 4:22-cv-7074-JST
16   Representative Plaintiffs,
                                              DEFENDANTS GITHUB AND
17   v.   **MICROSOFT'S NOTICE OF MOTIONS**
                                              AND MOTIONS TO DISMISS
18 GITHUB, INC., et al.,   **OPERATIVE COMPLAINT IN**
                                              CONSOLIDATED ACTIONS
19   Defendants.
                                              Date: May 4, 2023
20   Time: 2:00 p.m.
                                              Courtroom: 6, 2d Floor
21 AND CONSOLIDATED ACTION   Judge: Hon. Jon S. Tigar

22   Complaint Filed: December 7, 2022

23

24

25

26

27

28

1

TABLE OF CONTENTS

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

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

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

California Civil Code

Other Authorities

Federal Rules of Civil Procedure

1

NOTICE OF MOTIONS AND MOTIONS

2

TO PLAINTIFFS AND THEIR ATTORNEYS OF RECORD:

3

PLEASE TAKE NOTICE THAT on May 4, 2023 at 2:00 p.m., before the Honorable

4   Jon S. Tigar seated in Courtroom 6 of the United States Courthouse at Oakland, California, with

5   appearances to be made by Zoom videoconference unless otherwise ordered by the Court,

6   Defendants GitHub, Inc. ("GitHub") and Microsoft Corporation ("Microsoft") will, and hereby

7   do, move, pursuant to Federal Rules of Civil Procedure 9, 10, 12(b)(1) and 12(b)(6), to dismiss

8   the operative Complaint in these consolidated actions, filed as ECF No. 1 in the 22-cv-07074

9   action (the "Complaint" or "Compl."), in its entirety as to both GitHub and Microsoft. *See* ECF

10   No. 47 (ordering consolidation and designation of the operative Complaint).

11   The grounds for the Motions are as follows. First, the Complaint should be dismissed in

12   its entirety pursuant to Rules 10 and 12(b)(1) because Plaintiffs fail to identify themselves and fail

13   to allege that they suffered any actual or threatened injury at the hands of either GitHub or

14   Microsoft sufficient to give rise to a concrete controversy to be adjudicated by this Court.

15   Second, the Complaint should be dismissed in its entirety pursuant to Rule 12(b)(6) (and Rule 9

16   with respect to Count IV for Fraud) because each and every one of the numerous claims for relief

17   interposed against GitHub and Microsoft are either (a) legally foreclosed, (b) deficient because

18   factual allegations required to support necessary elements are missing or implausible, or (c) both.

19

20

21

22

23

24

25

26

27

28

GITHUB AND MICROSOFT'S MTD
No. 4:22-CV-6823-JST

1

MEMORANDUM OF POINTS AND AUTHORITIES

2

INTRODUCTION AND SUMMARY OF ISSUES

3   Since the advent of open source software, developers have been free—indeed,

4   encouraged—to use, study, change, and share source code for the broader good. It's an

5   extraordinary idea, one that has fostered an immense body of public knowledge, ever-evolving

6   and ever-available for the next generation of developers to build, collaborate, and progress. The

7   open source model is essential to collaborative and communal software development. GitHub

8   was founded on the basis of these ideals, and when Microsoft invested billions to acquire GitHub

9   in 2018, it cemented its commitment to them. The transformative technology at issue in this case,

10 Copilot, reflects GitHub and Microsoft's ongoing dedication and commitment to this profound

11 human project. Copilot is a coding assistant tool that crystallizes the knowledge gained from

12 billions of lines of public code, harnessing the collective power of open source software and

13 putting it at every developer's fingertips.

14 Two anonymous detractors bring this lawsuit, claiming "software piracy on an

15 unprecedented scale," and seeking penalties in excess of $9 billion. Despite that rhetoric, they do

16 not advance a copyright infringement claim at all—doubtless an attempt to evade the limitations

17 on the scope of software copyright and the progress-protective doctrine of fair use. Copilot

18 withdraws nothing from the body of open source code available to the public. Rather, Copilot

19 helps developers write code by generating suggestions based on what it has learned from the

20 entire body of knowledge gleaned from public code. In so doing, Copilot advances the very

21 values of learning, understanding, and collaboration that animate the open source ethic. With

22 their demand for an injunction and a multi-billion dollar windfall in connection with software that

23 they willingly share as open source, it is Plaintiffs who seek to undermine those open source

24 principles and to stop significant advancements in collaboration and progress.

25 Plaintiffs' Complaint fails on two intrinsic defects: lack of injury and lack of an otherwise

26 viable claim. Plaintiffs do not allege that this extraordinary new tool has harmed them in any

27 way. They do not explain how teaching Copilot about their code has taken anything from them.

28 They do not allege that Copilot has done anything improper with *their* code, nor that GitHub or

GitHub And Microsoft's MTD
No. 4:22-cv-6823-JST

1   Microsoft did anything improper with *their* personal information. They seem to say that Copilot

2   could theoretically suggest a snippet of code that matches something they have published, and do

3   so without giving them proper attribution. How that series of hypothetical events could harm

4   them is also unexplained. There is no case or controversy here—only an artificial lawsuit brought

5   by anonymous Plaintiffs built on a remote possibility that they will fail to be associated with an

6   AI-generated code snippet that could in theory be connected to copyrightable aspects of their

7   source code. The Complaint's failure to make out a case for actual injury to named people

8   requires dismissal under Rules 10 and 12(b)(1). Part I, *infra*.

9   Additionally, Plaintiffs do not state a claim. Their Complaint cycles through twelve

10   purported counts, some of which embrace multiple theories, in search of a hook for their abstract

11   grievance against Copilot. Despite their attempt to make a federal case out of it, none of these

12   counts plausibly establishes that training or using Copilot violates any legal right. Part II, *infra*.

13   **ALLEGATIONS OF THE OPERATIVE COMPLAINT**

14   **A.   OpenAI Develops A Generative AI Tool Called Codex.**

15   OpenAI is a nonprofit organization that develops machine learning models, also referred

16   to as "Artificial Intelligence." Compl. ¶¶ 2, 126, 128. Such models are typically trained through

17   exposure to a corpus of material called "training data." Compl. ¶ 79. The patterns discerned

18   from the set of training data become part of the model, which can then generate answers based

19   upon those patterns in response to user prompts. Compl. ¶ 81; *see* Compl. ¶ 52.

20   The model at issue in this case is called Codex. Codex is a generative AI model trained

21   on publicly available computer source code. Compl. ¶¶ 52, 135. The model embodies "inferred

22   … statistical patterns governing the structure of code," Compl. ¶ 52, which it has discerned from

23   the training data based on "a complex probabilistic process." Compl. ¶ 79. It is thus capable, in

24   response to a prompt, of "predic[ting] … the most likely [coding] solution." Compl. ¶ 79.

25   "[E]ssentially [it] returns the solution it has found in the most [coding] projects when those

26   projects are somehow weighted to adjust for whatever variables [the model] ha[s] identified as

27   relevant." Compl. ¶ 79.

28   Generative AI models are capable of "simulat[ing] human reasoning or inference,"

2

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

engaging in the same sort of pattern recognition, synthesis, and prediction we do. Compl. ¶ 2. AI models like Codex also enable statistical analysis and prediction vastly more powerful, efficient, and sensitive than what the human brain can accomplish. Compl. ¶ 81. At the same time, Codex "does not understand the meaning of code," Compl. ¶ 54, nor its "semantics and context the way humans do," Compl. ¶ 81. Codex thus offers both the ingenious and the mundane, a powerful tool of invention for humans who supply the insight to direct its range of performance.

B.   GitHub Offers Copilot, A Code Completion Tool Built On Codex.

GitHub Copilot is a programming assistant. Compl. at 8 n.4, ¶¶ 82-83. It "uses the OpenAI Codex to suggest code and entire functions in real-time" to software developers. Compl. ¶ 47. To use Copilot, a GitHub user installs it "as an extension to various code editors, including Microsoft's Visual Studio and VS Code." Compl. ¶ 67. "As the user types [code] into the editor," Copilot treats the user's input as a prompt, generating suggestions for code that may be appropriate for the developer's purposes. Compl. ¶ 67. Copilot is a subscription tool available to GitHub users for $10 per month or $100 per year. Compl. ¶ 8.

The version of Codex that powers Copilot was trained on billions of lines of code that GitHub users stored in public GitHub repositories. *See* Compl. ¶ 143. When GitHub users put their code on GitHub, they choose whether to make the code repositories private or public. Compl. ¶ 119. Users who set their repositories "to be viewed publicly … grant each User of GitHub a nonexclusive, worldwide license to use, display, and perform Your Content through the GitHub Service and to reproduce Your Content solely on GitHub as permitted through GitHub's functionality." Compl. Ex. 1 at 28 (GitHub Terms of Service ("TOS") at 7). Every user agrees to GitHub's TOS, which include a "License Grant" to GitHub to "store, archive, parse, and display … and make incidental copies" as well as "parse it into a search index or otherwise analyze it" and "share" the content in public repositories with other users. Compl. Ex. 1 at 27-28 (GitHub TOS at 6-7). And users can also select from a range of preset open source licenses to apply to the code published in their various GitHub repositories, apply their own individual licenses, or select none at all. Compl. ¶ 34 n.4 & Appx. A.

Any GitHub user thus appreciates that code placed in a public repository is genuinely

1   public. Anyone is free to examine, learn from, and understand that code, as well as repurpose it

2   in various ways. And, consistent with this open source ethic, neither GitHub's TOS nor any of

3   the common open source licenses prohibit either humans or computers from reading and learning

4   from publicly available code. *See* Compl. ¶ 34 n.4 & Appx. A.

5   **C.   Plaintiffs Sue Based On An Attribution Theory.**

6   Plaintiffs are two anonymous GitHub users. Compl. ¶¶ 19-20. They both claim to have

7   (at an unspecified time) "published Licensed Materials they owned a copyright interest in to at

8   least one GitHub repository under one of the Suggested licenses." Compl. ¶¶ 19-20. But they do

9   not allege, either expressly or on information and belief, that their Licensed Materials were used

10   to train Codex or Copilot. The most charitable reading of the Complaint is that because their

11   Licensed Materials were in public repositories, and since public repositories were used to train

12   Codex and Copilot, Compl. ¶ 82, they believe their Licensed Materials were used to do so. The

13   Complaint also is not clear on whether or to what extent the training of Copilot forms the basis of

14   the Plaintiffs' claims for relief.

15   The crux of the Complaint, instead, is Copilot's suggestions, which Plaintiffs term

16   "Output." Plaintiffs allege that these suggestions may sometimes match snippets of code used to

17   train Copilot, but without providing information like authorship or licensing status. According to

18   Plaintiffs, Copilot's "Output is often a near-identical reproduction of code from the training data,"

19   Compl. ¶ 46, and Copilot "has not been trained to provide Attribution." Compl. ¶ 56. On this

20   basis, Plaintiffs allege that "Defendants stripped Plaintiffs' and the Class's attribution, copyright

21   notice, and license terms from their code." Compl. ¶¶ 10-11, 78-81. Beyond that conclusory

22   allegation, however, Plaintiffs do not connect their own code to any such Outputs. They identify

23   no Output that has matched any of Plaintiffs' Licensed Materials. They identify no prompt that

24   might produce such a match. They allege only a few examples of code Outputs matching *other*

25   authors' code that have nothing to do with Plaintiffs. Compl. ¶¶ 48-81. And, while they assert

26   that Copilot's suggestions will "often" match existing code, they point only to a study suggesting

27   that "about 1% of the time, a suggestion … may contain some code snippets longer than ~150

28   characters that matches" some preexisting code. Compl. ¶ 90. Plaintiffs' allegations concerning

4

1   Copilot do nothing to connect their own code to an Output.

2   The Complaint nominally advances a dozen claims for relief related to Codex and Copilot.

3   The first claim encompasses multiple types of alleged violations and both direct and secondary

4   theories of liability. Compl. ¶¶ 142-71 (Count I). Other claims specify a remedy with multiple

5   legal theories, or otherwise rely on multiple theories. Compl. ¶¶ 204-10 (Count VI); *id.* ¶¶ 211-

6   14 (Count VII). Plaintiffs also tack on a theory of "civil conspiracy" covering most of the claims.

7   Compl. ¶¶ 240-44.

8   But this long and complex Complaint speaks loudest with what it doesn't say. Plaintiffs

9   fail to identify any of the "Licensed Materials" they allegedly placed in a GitHub public

10   repository that reflect purported "copyright interests," or to tell us anything at all about those

11   materials. The Complaint nowhere identifies any copyrighted work owned by either of the

12   Plaintiffs, or any registration of such work. The Complaint fails to identify any use of their

13   Licensed Materials. Although the case is supposedly about "software piracy on an unprecedented

14   scale," Plaintiffs make no copyright infringement claim. And, Plaintiffs identify no personal

15   identifying information that they stored in their public repositories on GitHub, or say how it was

16   allegedly exposed by Codex or Copilot.

17   <div align="center">**ARGUMENT**</div>

18   **I.   PLAINTIFFS LACK ARTICLE III STANDING AND THEREFORE SUBJECT**
       MATTER JURISDICTION BECAUSE THEY HAVE NOT ALLEGED ACTUAL
19   **OR THREATENED INJURY.**

20   GitHub and Microsoft move to dismiss under Rule 12(b)(1) because Plaintiffs fail to plead

21   the actual or threatened injury necessary for Article III standing. If a plaintiff lacks Article III

22   standing, the federal court lacks subject matter jurisdiction and the suit must be dismissed.

23   *Warren v. Fox Family Worldwide, Inc.*, 328 F.3d 1136, 1140 (9th Cir. 2003). "'[A]t the pleading

24   stage, the plaintiff must clearly allege facts demonstrating each element of standing.'" *McGee v.*

25   *S-L Snacks Nat'l*, 982 F.3d 700, 705 (9th Cir. 2020) (cleaned up). One such element "plaintiffs

26   must demonstrate [is] that they suffered a concrete harm." *TransUnion LLC v. Ramirez*, 141

27   S. Ct. 2190, 2200 (2021). This "injury in fact" requirement demands "more than an injury to a

28   cognizable interest. It requires that the part[ies] seeking review be [themselves] among the

1   injured." *Lujan v. Defenders of Wildlife*, 504 U.S. 555, 563 (1992). A plaintiff must therefore

2   allege an injury "that affects the plaintiff in a personal and individual way." *McGee*, 982 F.3d at

3   705 (cleaned up). And it must do so "for each claim … and for each form of relief that is

4   sought." *Davis v. Federal Election Comm'n*, 554 U.S. 724, 734 (2008) (cleaned up).

5   The two anonymous Plaintiffs in this action, J. Does 3 and 4, do not meet this standard as

6   to either of the Complaint's inconsistent theories of harm. The first theory is that Plaintiffs

7   should have been identified in connection with their code (lack of attribution), and the second is

8   that they should *not* have been identified (privacy). Yet the Complaint lacks allegations sufficient

9   to demonstrate that Plaintiffs were or will be injured in either respect.

10   **A.   Plaintiffs' Lack-Of-Attribution Theory Is Insufficient To Confer Standing.**

11   Both of the Plaintiffs allege that they are GitHub users who have published unspecified

12   source code subject to unspecified open source licenses at an unspecified time. *See* Compl.

13   ¶¶ 19-20. For all their references to "copyright interests" and a "Brave New World of Software

14   Piracy," neither of the Plaintiffs allege ownership or infringement of any copyrighted (let alone

15   registered, *see* 17 U.S.C. § 411) work. They do not even *identify* a copyrighted work. Plaintiffs

16   thus allege no invasion of their copyright interests—an allegation that would run headlong into

17   the doctrine of fair use. *See Google LLC v. Oracle Am., Inc.*, 141 S. Ct. 1183 (2021); *Author's*

18   *Guild v. Google, Inc.* 804 F.3d 202 (2d Cir. 2015); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d

19   1146 (9th Cir. 2007); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003); *Sega Enters. Ltd.*

20   *v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992).

21   Nor do Plaintiffs identify any harm that has come from the bare use of the contents of

22   public repositories to train Codex or Copilot. Plaintiffs admittedly chose to make their source

23   code freely available for inspection by anyone. Plaintiffs assert no legal interest that would

24   entitle them to restrict the study by human or machine of freely available code. Plaintiffs do not

25   explain how tangible harm to them results from an AI model's study of such code for training

26   purposes. And the lack of such an allegation is no surprise, since the First Amendment and fair

27   use generally protect such study. The open source principles embodied in common license

28   agreements embrace learning and understanding from published code.

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   Plaintiffs instead conjure a theory of harm related not to copyright interests or training, but

2   to lack of attribution. The apparent contention is that Copilot may generate suggested snippets of

3   code that match snippets in existing GitHub projects, without providing credit. As explained

4   below in connection with Microsoft and GitHub's Rule 12(b)(6) motion, Plaintiffs fail to state a

5   viable claim based on this theory. But even assuming such a claim is conceivable in the abstract,

6   there is no factual allegation in the Complaint suggesting that *these Plaintiffs* have suffered injury

7   under their snippet-without-credit theory. They allege a few examples of a Copilot suggestion

8   matching *someone else's* code. Compl. ¶¶ 66-76. They perform a spurious back-of-the-envelope

9   calculation suggesting 12,000 users whose code might be matched by a Copilot suggestion.

10   Compl. ¶ 91. Nothing in the Complaint, however, suggests that either of the Plaintiffs have been

11   or will be among those users. That leaves no one in this case claiming that Copilot has actually

12   caused them any harm at all—let alone pleading a *legal claim* based on that harm.

13   Plaintiffs also fail to allege facts supporting a reasonable inference that their code

14   plausibly would become a Copilot Output. What code do they claim an interest in? What

15   problem does that code solve? Where and how frequently has their published code been

16   replicated by others such that Copilot would identify it as "the most likely solution to a given

17   prompt," Compl. ¶ 79? What prompt could potentially generate a match, and why is it likely that

18   a user would enter this prompt? Without these types of allegations, Plaintiffs allege at best the

19   sort of "conjectural or hypothetical" injury that cannot confer standing. *Lujan*, 504 U.S. at 650.

20   Indeed, Plaintiffs' failure to identify *themselves* in violation of Rule 10 underscores the

21   wholly speculative nature of the case. The "use of fictitious names runs afoul of the public's

22   common law right of access to judicial proceedings" and is counter to "Rule 10(a)'s command

23   that the title of every complaint 'include the names of all the parties.'" *Does v. Advanced Textile*

24   *Corp.*, 214 F.3d 1058, 1067 (9th Cir. 2000) (quoting Fed. R. Civ. P. 10(a)).[*] If what Plaintiffs

25   _____

[*] Plaintiffs have not obtained leave to proceed without identifying themselves, nor do they allege
26   facts that would meet one or more of three scenarios justifying a use of pseudonyms: (1) where
the anonymous party faces retaliation, (2) where the litigation concerns a "sensitive and highly
27   personal" matter, and (3) where the anonymous party admits an intention to engage in illegal
conduct. *Advanced Textile*, 214 F.3d. at 1068; *see generally Doe v. Kamehameha Sch./Bernice*
28   *Pauahi Bishop Est.*, 596 F.3d 1036, 1043-44 (9th Cir. 2010) (noting risk of "physical harm" or

1 truly sought was attribution for their code, one would expect them to identify themselves and that

2 code. Instead, they offer only abstraction.

3 **B.   Plaintiffs Also Do Not Allege Privacy-Based Injury.**

4 Even as Plaintiffs hypothesize harm from failing to be identified, they also allege their

5 privacy is injured through misuse of their personal identifying information ("PII"). Compl.

6 ¶¶ 225-39. This second and inconsistent theory of injury, like the first, fails to demonstrate any

7 actual or threatened harm because Plaintiffs have alleged no details concerning collection or

8 mishandling of any of their PII. They have not identified the type of information at issue, any

9 specific information of their own, how and from where any Defendant allegedly collected that

10 information, or how that information was mishandled.

11 Plaintiffs therefore cannot allege a privacy-based injury, like invasion of privacy, loss of

12 control over private information, or disclosure of private information. *Cf. TransUnion*, 141 S. Ct.

13 at 2204 (identifying types of harms supporting standing). Indeed, any such injury appears

14 *implausible* based on what is alleged: if Plaintiffs are claiming that their PII was collected from

15 *public* repositories on GitHub, it is difficult to see how collection of that information could

16 possibly effect a *privacy* harm. In any event, Plaintiffs' failure to allege privacy-based injury

17 requires dismissal. *E.g.*, *I.C. v. Zynga, Inc.*, 600 F. Supp. 3d 1034, 1048-50 (N.D. Cal. 2022).

18 **II.   PLAINTIFFS FAIL TO STATE A DMCA CLAIM.**

19 Apart from failing to establish standing, Plaintiffs also do not state a claim. To survive a

20 motion to dismiss under Rule 12(b)(6), a complaint must contain factual matter that, when

21 accepted as true, states a claim that is plausible on its face. *Ashcroft v. Iqbal*, 556 U.S. 662, 678

22 (2009). If a claim lacks a cognizable legal theory or lacks sufficient facts to plausibly support a

23 cognizable legal theory, dismissal is warranted. *Thunder Power New Energy Vehicle Dev. Co. v.*

24 *Byton N. Am. Corp.*, 340 F. Supp. 3d 922, 925 (N.D. Cal. 2018).

25 Count I alleges violations of 17 U.S.C. § 1202, a part of the Digital Millennium Copyright

26 Act ("DMCA"). The DMCA was enacted to balance the rights of users with those of copyright

27 _____

   "severe threats of deportation, arrest, and imprisonment" as justifying anonymity); *United States*
28 *v. Doe*, 655 F.2d 920, 922 n.1 (9th Cir. 1980) (pseudonyms permitted in cases concerning
   abortion and contraceptive bans).

owners in light of new technologies. *See generally United States v. Elcom Ltd.*, 203 F. Supp. 2d 1111 (N.D. Cal. 2002). Section 1202 is directed to the potential for bad acts that facilitate mass infringement in the electronic age. It protects the "[i]ntegrity of copyright management information" ("CMI")—that is, certain categories of "information conveyed in connection with copies … of a work," like titles or authorship information. 17 U.S.C. § 1202(c). Section 1202 proscribes the provision or distribution of false CMI, intentional removal of CMI, alteration of CMI, and distribution of works with removed or altered CMI. Consistent with its focus on thwarting piracy, each subsection contains a double-scienter requirement. *See id.* A defendant is liable only if it had both the requisite scienter with respect to the violative conduct (i.e., provision of false CMI, removal of CMI, etc.) *and* some degree of knowledge that this conduct would "induce, enable, facilitate, or conceal infringement." *Id.* § 1202(a)-(b). Neutral technology like Copilot cannot satisfy that standard. Part A, *infra*. Even were that possible as a theoretical legal matter, Count I lumps far too much together to provide each of GitHub and Microsoft fair notice of who is alleged to have done what to violate § 1202. Part B, *infra*. Plaintiffs' purported claims under each subsection fail for multiple additional reasons. Sections C-D, *infra*.

A.   Plaintiffs Cannot Allege The Required Likelihood Of Infringement.

All of the § 1202 claims fail because Plaintiffs do not allege any likelihood that Copilot will result in copyright infringement of any of their works, as required to state a claim under *Stevens v. Corelogic, Inc.*, 899 F.3d 666, 673 (9th Cir. 2018). To plead a claim under § 1202(a) pertaining to false CMI, a plaintiff must allege that the defendant acted "with intent to induce, enable, facilitate, or conceal infringement." A claim under § 1202(b) regarding altered or removed CMI similarly requires that a defendant either "have knowledge" or have "reasonable grounds to know" of the same result. Of necessity, then, the tampering with CMI contemplated by § 1202 must be part of a scheme involving purposeful infringement.

That line is critical in a case like this one, where a plaintiff attacks a dual-use technology—a system that is not aimed at infringement, but may nevertheless touch or concern CMI in some way. Whether or not they advance copyright infringement claims, Plaintiffs must

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   still explain in the context of their § 1202 claims how infringement is occurring or is likely to

2   occur. They cannot. At most, Plaintiffs allege Copilot's suggestions will match snippets of

3   existing code merely 1% of the time. Compl. ¶ 90. And even within that miniscule percentage,

4   they do not try to allege *copyright infringement*—a virtual impossibility anyway given the various

5   copyright-based obstacles (originality, fair use, etc.) that would preclude a claim. Plaintiffs do

6   not allege a likelihood of infringement, let alone *purposeful* infringement.

7   *Stevens* addressed this scenario in a case involving § 1202(b)'s "reasonable grounds to

8   know" requirement—the lowest of § 1202's scienter requirements. Photographers challenged the

9   defendant's software program, which "downsampled" (i.e., condensed the size of) images for use

10 in electronic real estate listings, and in the process failed to preserve the CMI metadata from the

11 plaintiffs' photographs. 899 F.3d at 671. The plaintiffs nonetheless argued that they satisfied the

12 "reasonable grounds to know" requirement because when CMI is removed, "one method of

13 identifying an infringing photograph has been impaired, [such that] someone *might* be able to use

14 their photographs undetected." *Id.* at 673.

15   The Ninth Circuit rejected the argument. Although a § 1202 plaintiff "need not show that

16 any specific infringement has already occurred," it still "must make an affirmative showing …

17 that the defendant was aware or had reasonable grounds to be aware of the probable future impact

18 of its actions." *Id.* at 674. And this requires a showing "from which one can infer that future

19 infringement is *likely* … to occur as a result of the removal or alteration of CMI." *Id.* at 675

20 (emphasis added). The rule makes good sense: If infringement is merely a "general possibility,"

21 *id.* at 673, there is no plausible inference that the defendant's conduct was done with the requisite

22 mental state. *See also Kelly v. Arriba Soft Corp.*, 77 F. Supp. 2d 1116, 1122 (C.D. Cal. 1999)

23 (rejecting DMCA claim based on "unintended side effect" of web crawler's operation), *rev'd on*

24 *other grounds*, 336 F.3d 811 (9th Cir. 2003). Courts have dismissed § 1202 claims at the

25 pleading stage for failing to satisfy this requirement. *See Harrington v. Pinterest, Inc.*, No. 20-cv-

26 05290, 2021 WL 4033031, at *6 (N.D. Cal. Sept. 3, 2021); *Philpot v. Alternet Media, Inc.*, No.

27 18-cv-4479, 2018 WL 6267876, at *5 (N.D. Cal. Nov. 30, 2018).

28   Much as in *Stevens*, Plaintiffs here challenge technology whose purpose is unrelated to

1   infringement or CMI, but that incidentally may not preserve CMI. Plaintiffs allege that Copilot

2   "has not been trained to provide Attribution," Compl. ¶ 56, and that, "[o]n information and belief,

3   Defendants could have trained Copilot to include attribution, copyright notices, and license

4   terms." Compl. ¶ 161. Plaintiffs thus allege that Copilot is CMI-agnostic. Affirmative steps

5   would have to be taken to modify the technology to account for CMI. That is the precise scenario

6   in which *Stevens* demands a showing that the technology's incidental effects with respect to CMI

7   have a likelihood of causing, facilitating, or concealing infringement.

8   To overcome *Stevens*, Plaintiffs would need to allege facts showing that the technology at

9   issue is likely to result in widespread infringement. Not only do Plaintiffs fail to allege such

10   facts, but it is obvious from the Complaint that Copilot is not such a machine. Plaintiffs allege

11   that suggestions match training data very rarely. Even assuming that were infringement, there is

12   no obligation in copyright law to modify technology to avoid all possibility of infringement. *See*

13   *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984). Finally, and most

14   importantly, it makes no sense to assume that any matching Outputs that might occur constitute

15   copyright infringement. Even when a snippet matches training data, there are a host of additional

16   hurdles to the notion that any particular Output would result in copyright infringement: the

17   copyrightability of such material in light of doctrines of merger and scènes à faire, *see Apple*

18   *Computer, Inc. v. Microsoft Corp.*, 35 F.3d 1435, 1444-45 (9th Cir. 1994); the requirement of

19   "virtually identical" copying of thinly protected elements, *see id.* at 1444; the scope of

20   authorization through licensing, *compare Sun Microsystems, Inc. v. Microsoft Corp.*, 188 F.3d

21   1115 (9th Cir. 1999), *with Jacobsen v. Katzer*, 535 F.3d 1373 (Fed. Cir. 2008); and fair use, *see*

22   *Google LLC v. Oracle Am., Inc.*, 141 S. Ct. 1183 (2021). Given these hurdles, Copilot is simply

23   not the type of technology that could ever satisfy the requirements of *Stevens*.

24   Similarly, just as in *Stevens*, Plaintiffs fail to offer any "specific [allegations] that removal

25   of CMI … will impair their policing of infringement." 899 F.3d at 675. They do not, for

26   example, "aver" that they "have ever used CMI … to prevent or detect copyright infringement."

27   *Id.* And just as "a party intent on using a copyrighted photograph undetected can itself remove

28   any CMI metadata," *id.* at 676, a party intent on using copyrighted code without reproducing

11

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   licensing information could readily do so. So, "one cannot plausibly say that removal [by

2   Copilot] 'will'" make it easier to infringe Plaintiffs' works. *Id.*

3   It is Plaintiffs' prerogative to not assert a claim for copyright infringement. But *Stevens*

4   requires plausible allegations that Copilot's operation with respect to CMI is likely to have that

5   result. Plaintiffs make no such allegations, and the nature of Copilot precludes them from doing

6   so. Count I should be dismissed with prejudice.

7   **B.   Count I Impermissibly Lumps Together Multiple Defendants, Claims, And**
         Theories Of Liability.
8

9   Plaintiffs cannot get around *Stevens*, but in any event the "failure to allege what role each

10 Defendant played in the alleged harm makes it exceedingly difficult, if not impossible, for

11 individual Defendants to respond to [the] allegations." *In re iPhone Application Litig.*, No. 11-

12 MD-02250, 2011 WL 4403963, at *8 (N.D. Cal. Sept. 20, 2011). Such a pleading therefore does

13 not satisfy Rule 8(a). *Id.*; *see Destfino v. Reiswig*, 630 F.3d 952, 958-59 (9th Cir. 2011)

14 (affirming dismissal of "shotgun pleading").

15   Count I is a textbook example of this problem. It is alleged against "All Defendants."

16 Over the course of a nearly 30-paragraph thicket of assertions, Plaintiffs appear to claim violation

17 of all five of the statute's prohibitions. Compl. ¶¶ 142-71. They also purport, by parenthetical

18 under the title of "COUNT I," to advance "(Direct, Vicarious, and Contributory)" theories of

19 liability. *Id.* at 34. No allegation is specific to either GitHub or Microsoft. Nothing elsewhere in

20 the Complaint explains what any particular Defendant is supposed to have done.

21   This defect is worst with respect to Microsoft. Plaintiffs allege that Microsoft owns

22 GitHub and is an investor in OpenAI. Compl. ¶¶ 5-7. But it is basic corporate law that absent

23 some disregard of corporate formalities, Microsoft cannot be held liable as a corporate parent or

24 shareholder. *See United States v. Bestfoods*, 524 U.S. 51, 61 (1998). As for Microsoft's

25 subsidiary GitHub, there is no explanation of what GitHub supposedly did with respect to the

26 alteration or removal of CMI. And since there is no allegation suggesting that GitHub or

27 Microsoft engaged in primary conduct involving CMI, Plaintiffs presumably mean to sweep them

28 in with the "vicarious" and "contributory" labels. But § 1202 does not provide for such theories

1   of liability on its face, and it is doubtful that the Ninth Circuit would find § 1202's double-

2   scienter requirement compatible with theories of secondary liability. *Cf. Petroliam Nasional*

3   *Berhad v. GoDaddy.com, Inc.*, 737 F.3d 546 (9th Cir. 2013) (refusing extra-statutory contributory

4   liability theory under anti-cybersquatting statute in part because of "bad faith" requirement).

5   Plaintiffs are claiming entitlement to statutory damages of over $9 billion for their § 1202

6   claims. Compl. at 52 n.42. Defendants cannot properly address these claims without knowing

7   what Plaintiffs say each Defendant did to violate the statute.

8   **C.   Plaintiffs Have Not Plausibly Alleged That GitHub Or Microsoft
               Intentionally Or Knowingly Removes Or Alters CMI Under § 1202(b).**
9

10   Plaintiffs' claims under § 1202(b)(1) and (3) fail for an independent reason. Section

11   1202(b)(1) prohibits a defendant from "intentionally remov[ing] or alter[ing] [CMI]," *see* Compl.

12   ¶ 148, while § 1202(b)(3) bars "distribut[ing] … copies of works … knowing that [CMI] has been

13   removed or altered," *see* Compl. ¶ 154. These provisions require the active step of removal or

14   alteration of CMI in connection with an identical copy of the work. But all the Complaint alleges

15   is the passive non-inclusion of CMI with a proffered snippet that may incidentally match a

16   portion of a work. That is not enough.

17   This result flows directly from § 1202's text. The terms "remove" and "alter" both

18   connote active conduct. To remove is "to move by lifting, pushing aside, or taking away or off"

19   or "to get rid of." Webster's Third New International Dictionary 1921 (1993). To alter is "to

20   cause to become different." *Id*. at 63. As explained above, CMI is defined as information

21   "conveyed in connection *with copies* … of a work"—that is, in or directly connected with "the

22   body of or the actual work itself." *See Mills v. Netflix, Inc.*, No. CV 19-7618, 2020 WL 548558,

23   at *3 (C.D. Cal. Feb. 3, 2020). To remove or alter CMI, then, is to engage in active conduct to

24   get rid of or change CMI that would otherwise naturally accompany a copy of the work from

25   which the CMI is allegedly removed or altered. *See Falkner v. Gen. Motors LLC*, 393

26   F. Supp. 3d 927, 938-39 (C.D. Cal. 2018).

27   Courts have consistently held this line. They have rejected § 1202(b) claims based on

28   mere "framing" of a photograph in a way that does not include CMI, *id.*; excerpting lecture notes

13   GitHub and Microsoft's MTD
                                        No. 4:22-cv-6823-JST

1 and study questions from textbooks without reproducing CMI, *Faulkner Press, L.L.C. v. Class*

2 *Notes, L.L.C.*, 756 F. Supp. 2d 1352, 1356, 1359 (N.D. Fla. 2010); copying "aspects" of

3 architectural works but "omitting" the plaintiff's CMI, *Design Basics, LLC v. WK Olson*

4 *Architects, Inc.*, No. 17 C 7432, 2019 WL 527535, at *5 (N.D. Ill. Feb. 11, 2019); *Frost-Tsuji*

5 *Architects v. Highway Inn, Inc.*, No. CIV. 13-00496 SOM, 2015 WL 263556, at *3 (D. Haw. Jan.

6 21, 2015), *aff'd*, 700 F. App'x 674 (9th Cir. 2017); or incorporating the underlying content from

7 the original copy into some different form or distinct work without CMI, *Kelly*, 77 F. Supp. 2d at

8 1122 (thumbnail versions of images). Courts have recognized § 1202(b) claims only for active

9 conduct in connection with identical copies of a work—like a defendant cropping a "gutter

10 credit" from an otherwise identical photo, *e.g.*, *Murphy v. Millennium Radio Grp. LLC*, 650 F.3d

11 295 (3d Cir. 2011), or deleting CMI *from within* identical content that was excerpted. *E.g.*,

12 *Bounce Exchange, Inc. v. Zeus Enter. Ltd.*, No. 15cv3268, 2015 WL 8579023, at *2-3 (S.D.N.Y.

13 Dec. 9, 2015) (CMI actively removed from lines of software code).

14   Plaintiffs allege only that Copilot "has not been trained" to actively identify and reproduce

15 any CMI that may or may not be present. Compl. ¶ 56. And rather than alleging removal of CMI

16 connected with an identical copy of any work, their claims are based on mere code suggestions

17 that *may* match small snippets of code contained in other works. Indeed, as Plaintiffs

18 acknowledge, Copilot's suggestions are not copied from any particular source, but generated

19 based on patterns it has "seen *the most*," Compl. ¶ 73. Plaintiffs therefore cannot state a claim

20 that CMI is actively removed from any work.

21   **D.**   **Plaintiffs Have Not Plausibly Alleged That GitHub Or Microsoft Provides,**

22   **Distributes, Or Imports For Distribution Any False CMI.**

23   In addition to the failure to plead likely copyright infringement, Section B, *supra*,

24 Plaintiffs' claims under § 1202(a)(1)-(2) and (b)(2) also fail for an independent reason. Those

25 provisions bar a defendant from "provid[ing] [CMI] that is false," § 1202(a)(1), *see* Compl.

26 ¶ 158; "distribut[ing] or import[ing] for distribution [CMI] that is false," § 1202(a)(2), Compl.

27 ¶ 158; and "distribut[ing] or import[ing] for distribution [CMI] knowing that the [CMI] has been

28 removed or altered," § 1202(b)(2), Compl. ¶ 153. But Plaintiffs fail to plausibly allege that

14

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   GitHub or Microsoft provide, distribute, or import for distribution any CMI at all.

2   Plaintiffs' only allegation in this regard is that "Defendants have a business practice of

3   asserting and/or implying that Copilot is the author of the Licensed Materials." Compl. ¶ 158.

4   Plaintiffs provide no instance of such an assertion or implication, either as to works users place in

5   public repositories or Copilot's suggestions. And GitHub's Terms of Service ("TOS"), attached

6   to the Complaint, directly refute Plaintiffs' allegation: "GitHub does not claim any rights in

7   [Copilot] Suggestions, and you retain ownership of and responsibility for Your Code, including

8   Suggestions you include in Your Code." Compl. Ex. 1 at 38 (Copilot Additional Terms at 1).

9   In any event, Plaintiffs do not allege any distribution of CMI by GitHub or Microsoft *at*

10   *all*, let alone false CMI. Section 1202(c) defines CMI as "information conveyed *in connection*

11   *with copies* … of a work." (emphasis added). Courts have construed the "in connection with

12   copies" language to require a specific, close nexus between the information conveyed and the

13   copy of the work. For example, a generic copyright notice on a webpage that was not "located on

14   or next to [the plaintiff's] photos" lacked sufficient connection. *SellPoolSuppliesOnline.com,*

15   *LLC v. Ugly Pools Ariz., Inc.*, 804 F. App'x 668, 670-71 (9th Cir. 2020); *see Logan v. Meta*

16   *Platforms, Inc.*, No. 22-cv-01847, 2022 WL 14813836, at *8 (N.D. Cal. Oct. 25, 2022). The

17   Complaint identifies no information provided anywhere, let alone in connection with a copy of

18   any work. As for their suggestion that GitHub or Microsoft "imply[]" authorship, implications

19   are not "information" that can be provided, distributed, or imported for distribution. They are

20   therefore not CMI at all. And even if they were, Plaintiffs point to no implication of authorship

21   made in connection with a copy of a work. Plaintiffs' claims under § 1202(a) and (b)(2) should

22   therefore be dismissed.

23   **III.   PLAINTIFFS FAIL TO STATE A BREACH-OF-LICENSE CLAIM.**

24   "Under California law, to state a claim for breach of contract a plaintiff must plead the

25   contract, plaintiffs' performance (or excuse for nonperformance), defendant's breach, and damage

26   to plaintiff therefrom." *Low v. LinkedIn Corp.*, 900 F. Supp. 2d 1010, 1028 (N.D. Cal. 2012)

27   (quotation marks omitted). Count II fails this test.

28

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   **A.**   **Plaintiffs Fail To Specify The Contract Provisions Allegedly Breached.**

2   Plaintiffs' breach of contract claim fails at the outset because the Complaint does not

3   identify the contract(s) or provision(s) that each Defendant allegedly breached. *See Low*, 990 F.

4   Supp. 2d at 1028. One would ordinarily see a contract claim alleging "plaintiff and defendant

5   entered into X contract and then defendant violated paragraph Y of that contract by doing Z."

6   None of that exists here. Instead, all Plaintiffs assert is that "Defendants [lumped together]

7   violated the [unspecified] Licenses governing use of the [unidentified] Licensed Materials by

8   using them to train Copilot." Compl. at 10. There is no identification of the who and how of the

9   contractual relationships, no identification of the object(s) of any alleged contract, nor

10   identification of how any particular act by any particular Defendant violated any particular

11   contractual provision. Plaintiffs do not allege *how* training Copilot on public GitHub repositories

12   breaches any provision of any contract. This is not just procedural theater. It is important

13   *because none of the open source licenses attached to the Complaint appear to prohibit such*

14   *activity in any way*. On the contrary, the principles embodied in customary open source licenses

15   contemplate broad public rights to inspect, learn from, and build upon code.

16   Rule 8 requires at least a minimal effort in a contract claim to explain the breach. Count II

17   is defective in this respect, and the failure to offer any explanation is particularly egregious with

18   respect to Defendant Microsoft, which is not alleged to have done *anything* with Plaintiffs'

19   "Licensed Materials." Because Plaintiffs fail to plead the "specific provisions in the contract …

20   the defendant is said to have breached," *Young v. Facebook, Inc.*, 790 F. Supp. 2d 1110, 1117

21   (N.D. Cal. 2011), their claims should be dismissed.

22   **B.**   **To The Extent Count II Is Based On Training Copilot With Code Found In**
             Public Repositories, It Is Foreclosed By GitHub's Terms Of Service.
23

24   In all events, to the extent Plaintiffs' unexplained breach theory against GitHub is

25   predicated on training Copilot with code published in public GitHub repositories, Count II is

26   foreclosed by GitHub's Terms of Service. Compl. Ex. 1.

27   Consistent with open source licensing principles, GitHub's TOS expressly authorizes the

28   training of Copilot. The TOS informs users that "You own content you create, but you allow us

certain rights to it." Compl. Ex. 1 at 27 (TOS at 6). "These license grants apply to Your

Content" when uploaded to GitHub, notwithstanding any other license terms that might be

attached. *Id.* Every GitHub user chooses whether to allow their "repositories to be viewed

publicly." *Id.* No user is required to make a repository public. If they choose to do so, they

"grant [GitHub] … the right to store, archive, parse, and display" publicly posted content and

"make incidental copies, as necessary to provide the Service, including improving the Service

over time." *Id.* This includes the right to "copy" public material "to our database," "parse it into

a search index or otherwise analyze it on our servers," and "share it with other users." *Id.* at 27-

28 (TOS at 6-7). The "Service" includes all "the applications, software, products, and services

provided by GitHub," *id.* at 24 (TOS at 3), which includes Copilot. *See* Compl. ¶ 8; Compl. Ex. 1

at 24, 38 (TOS at 3, Copilot Additional Terms).

As Plaintiffs allege, Copilot functions by evaluating the code contained in GitHub's public

repositories, Compl. ¶¶ 82-83, in order to "infer[] statistical patterns governing the structure of

code." Compl. ¶ 52. That activity is squarely encompassed by the contractual authorization,

from every public code repository owner, for GitHub to "store, archive, parse, and display"

publicly posted content "as necessary to provide … the applications, software, products, and

services provided by GitHub" and to "analyze it on our servers." Compl. Ex. 1 at 27-28, 24 (TOS

at 6-7, 3). Even if the Complaint put forward any plausible theory of how inspection and analysis

of publicly posted material violated an open source license provision, Plaintiffs' acceptance of the

GitHub TOS forecloses a claim for breach of contract based on Copilot's training on the public

code. Dismissal is proper where a complaint alleges facts that demonstrate the claim is barred as

a matter of law. *Orchard Supply Hardware LLC v. Home Depot USA, Inc.*, 967 F. Supp. 2d

1347, 1354 (N.D. Cal. 2013). So it is here.

C.   Plaintiffs Have Not Plausibly Alleged Breach Based On An Output Theory.

Nor do Plaintiffs plausibly allege breach based on the use of Copilot. Plaintiffs allege that

"when providing Output, Copilot does not comply with [open source license] terms." Compl.

¶ 174. But Plaintiffs have not alleged that Copilot has provided any of *their* code subject to an

attribution requirement, nor plausibly pled that such a suggestion is likely. Nor do they explain

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   how, absent that, there is a breach by any of these Defendants. So even if Defendants knew

2   which provisions of which contracts were at issue, Plaintiffs have not plausibly alleged a breach

3   or plausibly alleged any damages to them—both essential elements of their claim. *Low*, 900 F.

4   Supp. 2d at 1028; *see also* Part I, *supra*.

5   Plaintiffs allege several examples of Copilot suggestions matching others' code. *See*

6   Compl. ¶¶ 66-76; Compl. ¶ 87 ("code from the game Quake III"); Compl. ¶ 88 (unspecified "code

7   that had been released under a license that allowed its use only for free games"); Compl. ¶ 89

8   ("code belonging to" a "Texas A&M computer-science professor"). None of these alleged

9   licensors appear to be named Plaintiffs (and Plaintiffs would be hard pressed to claim that

10 Microsoft—which owns Quake—is suing itself and its subsidiary in this lawsuit). Plaintiffs

11 cannot save their claims with surmise that "if Licensed Materials have been posted to a GitHub

12 public repository," they "can be reasonably certain [those materials are] sometimes returned to

13 users as Output." Compl. ¶ 82. "1% of the time," Compl. ¶ 90, is hardly likely—and in any

14 event, Plaintiffs fail to allege how even this 1% would breach any license.

15 **IV.   PLAINTIFFS' TORT AND UCL CLAIMS FAIL.**

16   Plaintiffs advance claims against GitHub and Microsoft under California law for tortious

17 interference with contract and economic advantage (Count III), against GitHub for fraud (Count

18 IV), and against GitHub for unlawful competition under the UCL (Counts VI and VII). Each is

19 preempted by the Copyright Act and deficient in other respects.

20   **A.   Plaintiffs' Tort And UCL Claims Are Preempted By The Copyright Act.**

21   Section 301 of the Copyright Act preempts "legal or equitable rights that are equivalent to

22 any of the exclusive rights within the general scope of copyright … and come within the subject

23 matter of copyright." 17 U.S.C. § 301(a). Courts evaluate copyright preemption under a "two-

24 part test," holding state law claims preempted where (1) "the 'subject matter' of the state law

25 claim falls within the subject matter of copyright as described in 17 U.S.C. §§ 102 and 103"; and

26 (2) "the rights asserted under state law are equivalent to the rights contained in 17 U.S.C. § 106."

27 *Maloney v. T3Media, Inc.*, 853 F.3d 1004, 1010 (9th Cir. 2017) (quotation marks omitted).

28   The subject matter requirement of the test is plainly satisfied. Plaintiffs' tortious

18

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   interference, fraud, and UCL claims all involve Plaintiffs' claimed copyright interests in

2   computer code, which constitutes a "literary work" under § 102. *See Google*, 141 S. Ct. at 1196.

3   Plaintiffs' tortious interference claim alleges that Defendants "improperly use[] Copilot to create

4   Derivative Works," Compl. ¶ 189, the subject matter of § 103. The fraud claim likewise is based

5   on the allegation that Defendants "sell[] … Licensed Materials as part of Copilot." Compl. ¶ 197.

6   And aspects of Plaintiffs' UCL claims based on non-privacy state law causes of action likewise

7   involve "code." Compl. ¶ 212(b). These claims therefore involve the "subject matter" of

8   copyright, a result that obtains even if the code is not actually protected by the Copyright Act.

9   *See Montz v. Pilgrim Films & Television, Inc.*, 649 F.3d 975, 979 (9th Cir. 2011) (en banc)

10   ("[T]he scope of the subject matter of copyright law is broader than the protections it affords.").

11   The second step is also met with respect to each claim—indeed, the claims quote nearly

12   verbatim from the Copyright Act's list of exclusive rights. As noted above, Plaintiffs' tortious

13   interference claim is based on the allegation that Defendants "create Derivative works," Compl.

14   ¶ 189, mirroring § 106(2)'s grant of the exclusive right "to prepare derivative works." And

15   Plaintiffs allege no other conduct or element of their tortious interference claim that would render

16   it "qualitatively different" from a claim of infringement of the right to create derivative works.

17   *Crafty Prods., Inc. v. Fuqing Sanxing Crafts Co.*, 839 F. App'x 95, 98-99 (9th Cir. 2020) (tortious

18   interference claim preempted); *Harper & Row Publishers, Inc. v. Nation Enters.*, 723 F.2d 195,

19   201 (2d Cir. 1983) (tortious interference claim preempted where "unauthorized publication is the

20   gravamen of the[] claim"), *rev'd on other grounds*, 471 U.S. 539 (1985).

21   Similarly, Plaintiffs' fraud claim alleges merely that GitHub stated that it would not sell or

22   distribute Plaintiffs' code, then "failed to honor [these] representations" by selling or distributing

23   Plaintiffs' code. Compl. ¶¶ 195-97. This claim is indistinguishable from one for infringement of

24   § 106(3)'s distribution right. Although fraud claims may escape preemption where they involved

25   intentional misrepresentations tangentially related to the subject matter of copyright, Plaintiffs

26   allege no misrepresentation at all—their claim is just an attempt to impose liability for

27   distributing code. This disguised copyright claim is preempted. *See Idema v. Dreamworks, Inc.*,

28   162 F. Supp. 2d 1129, 1191-92 (C.D. Cal. 2001); *Giddings v. Vison House Prod., Inc.*, No. CV

1   05-2963, 2007 WL 2274800, at *3 (D. Ariz. Aug. 7, 2007).

2   To the extent Plaintiffs' UCL claim is predicated on the above torts, it too is preempted.

3   *See Maloney*, 853 F.3d at 1019 (holding UCL claim preempted); *Kodadek v. MTV Networks, Inc.*,

4   152 F.3d 1209, 1213 (9th Cir. 1998) (same). And the same analysis requires dismissal of

5   Plaintiffs' invocation, under § 17200 and the common law, of the remedy of unjust enrichment,

6   which once again is based on the allegation that Defendants "used [Plaintiffs' code] to create

7   Derivative Works" and "res[old] [them]." Compl. ¶¶ 204-10.

8   **B.**   **Plaintiffs Do Not Plausibly Allege A Contract Or Business Expectancy, Nor Defendants' Knowledge Of One As Required For An Interference With**

9   **Economic Advantage Claim.**

10   Plaintiffs' tortious interference claim (Count III) against GitHub and Microsoft is deficient

11   in multiple other ways. "Tortious interference with contractual relations requires (1) the

12   existence of a valid contract between the plaintiff and a third party; (2) the defendant's knowledge

13   of that contract; (3) the defendant's intentional acts designed to induce a breach or disruption of

14   the contractual relationship; (4) actual breach or disruption of the contractual relationship; and (5)

15   resulting damage." *Ixchel Pharma, LLC v. Biogen, Inc.*, 9 Cal. 5th 1130, 1141 (2020) (quotation

16   marks omitted). The Complaint falls short on every element.

17   Although Plaintiffs label their claim one for tortious interference with *contract*, they

18   identify no contractual relationship between themselves and anyone else—much less a contract

19   that GitHub and Microsoft knew about. *See* Compl. ¶¶ 188-93; *see also Bed, Bath & Beyond of*

20   *La Jolla, Inc. v. La Jolla Vill. Square Venture Partners*, 52 Cal. App. 4th 867, 878 (1997)

21   (plaintiff must allege an existing, enforceable contract). While Plaintiffs allege that they

22   deposited materials in public GitHub repositories subject to open source licenses, they do not

23   allege that any third party ever entered into a contract with them in regard to those materials. Nor

24   do Plaintiffs plausibly allege that GitHub or Microsoft committed "intentional acts designed to

25   induce a breach or disruption" or "actual breach or disruption" with respect to an open source

26   license granted by Plaintiffs to any unidentified third party. *Ixchel*, 9 Cal. 5th at 1141. All they

27   allege is that "[a]t GitHub's upcoming yearly conference, GitHub Universe 2022, it will host a

28   presentation called 'How to compete with open source—and win.'" Compl. ¶ 190. Plaintiffs

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1   construct their claim on a tortured interpretation of this title, as though the presentation provides a

2   blueprint on how to compete *against* open source, when the more natural reading is that it is

3   about how to compete effectively *using* open source. But in all events, this is hardly an allegation

4   of a specific act to knowingly interfere with specific contracts. And in any event, "compet[ing]"

5   and "win[ning]" is not a tort. *See Ixchel*, 9 Cal. 5th at 1142. The Complaint therefore states no

6   claim for intentional interference with contract against GitHub or Microsoft.

7   The Complaint also fails to state a claim for the related tort of interference with

8   prospective economic advantage. Although "[t]ortious interference with prospective economic

9   advantage does not depend on the existence of a legally binding contract," it still requires a

10   plaintiff to "show that the defendant knowingly interfered with an economic relationship between

11   the plaintiff and some third party, which carries the probability of future economic benefit to the

12   plaintiff." *Ixchel*, 9 Cal. 5th at 1141 (cleaned up). Specific known relationships must be

13   identified. *Westside Ctr. Assocs. v. Safeway Stores 23, Inc.*, 42 Cal. App. 4th 507, 523-28 (1996);

14   *Rosen v. Uber Techs., Inc.*, 164 F. Supp. 3d 1165, 1178-79 (N.D. Cal. 2016).

15   Here, Plaintiffs fail to allege any specific relationship with a reasonably likely economic

16   benefit that GitHub and Microsoft knew about and intended to interfere with. The Complaint

17   merely asserts that "Defendants have wrongfully interfered with the business interests and

18   expectations of Plaintiffs and the Class by improperly using Copilot to create Derivative Works

19   that compete against OSC." Compl. ¶ 189. This conclusory "the-defendant-unlawfully-harmed-

20   me accusation," *Iqbal*, 556 U.S. at 678, does not identify the "economic relationship" supposedly

21   interfered with or plausibly allege the "probability of future economic benefit" claimed to be

22   disrupted. Nor does the Complaint allege "actual disruption" or specific "harm," two additional

23   elements of the claim. *Korea Supply Co. v. Lockheed Martin Corp.*, 29 Cal. 4th 1134, 1153

24   (2003). As a result, the Complaint does not state a claim for any economic interference tort.

25   **C.**   **Plaintiffs' Fraud Claim Against GitHub Is Barred By The Economic Loss
26   Rule And Fails Rule 9.**

27   "The elements of fraud … are (a) misrepresentation (false representation, concealment, or

28   nondisclosure); (b) knowledge of falsity (or 'scienter'); (c) intent to defraud, i.e., to induce

1 reliance; (d) justifiable reliance; and (e) resulting damage." *Lazar v. Superior Ct.*, 12 Cal. 4th

2 631, 638 (1996) (quoting 5 Witkin, Summary of Cal. Law (9th ed. 1988) Torts, § 6765, p. 778).

3 Federal Rule of Civil Procedure 9(b) further requires Plaintiffs to "state with particularity the

4 circumstances constituting fraud." This required "specificity includ[es] an account of the time,

5 place, and specific content of the false representations as well as the identities of the parties to the

6 misrepresentations." *Swartz v. KPMG LLP*, 476 F.3d 756, 764 (9th Cir. 2007) (cleaned up).

7   Plaintiffs assert that "in both its Terms of Service and its Privacy Statement, GitHub

8 promises not to sell Licensed Materials or anything else uploaded to or shared with GitHub."

9 Compl. ¶ 195. But Plaintiffs nowhere allege that GitHub made these statements with "knowledge

10 of falsity," "intent to defraud," or to "induce [unjustified] reliance." *Lazar*, 12 Cal. 4th at 638.

11 The Complaint states merely that Plaintiffs "relied upon [GitHub's] representations in choosing to

12 upload Licensed Materials to GitHub," in part because "GitHub has long held itself out as the best

13 place to host open source code repositories." Compl. ¶ 196. These vague assertions do not

14 suggest GitHub's intentions to commit fraud, particularly in light of the other provisions of the

15 TOS. *See supra* 3, 15-17. Nor, as discussed, have Plaintiffs alleged any "resulting damage."

16   Ultimately, then, the basis of Plaintiffs' claim is merely that GitHub "failed to honor its

17 representations." Compl. ¶ 197. That is a breach of contract claim, and thus barred by the

18 economic loss rule. "The economic loss rule" dictates that "purely economic loss due to

19 disappointed expectations" cannot be remedied in tort "unless [the Plaintiff] can demonstrate

20 harm above and beyond a broken contractual promise," therefore "prevent[ing] the law of

21 contract and the law of tort from dissolving one into the other." *Robinson Helicopter Co. v. Dana

22 Corp.*, 34 Cal. 4th 979, 988 (2004) (cleaned up). The fraud claim against GitHub is merely

23 derivative of alleged promises in the contractual TOS and the Complaint identifies no loss, let

24 alone one independent of the promises of the TOS. Plaintiffs' fraud claims should be dismissed.

25   **D.   Plaintiffs Lack Standing To Bring A UCL Unlawful Competition Claim And
          Fail To State A Claim.**

26

27   Plaintiffs duplicate their DMCA, tort, passing off, and privacy claims as purported

28 violations by GitHub of California's Unfair Competition Law, Cal. Bus. & Prof. Code §§ 17200,

1   *et seq.* (Count VII). This is dead on arrival. The "UCL's standing requirements" are even "more

2   stringent than the federal standing requirements." *Kwikset Corp. v. Super. Ct.*, 51 Cal. 4th 310,

3   324 (2011). A private plaintiff may only sue under the UCL if it has "suffered injury in fact and

4   has lost money or property." Cal. Bus. & Prof. Code § 17204. Standing for the UCL is

5   "substantially narrower than federal standing under Article III." *Kwikset*, 51 Cal. 4th at 324; *see*

6   *also McGee*, 982 F.3d at 705 n.3 (same). Here, because Plaintiffs lack Article III standing, *see*

7   Part I, *supra*, they certainly lack UCL standing. Plaintiffs do not plausibly allege the required

8   "lost money or property." Cal. Bus. & Prof. Code § 17204. They assert that "Plaintiffs and the

9   Class have suffered monetary damages as a result of GitHub's and OpenAI's conduct." Compl.

10   ¶ 213. But Plaintiffs do not allege any factual theory of *how* they suffered monetary loss as a

11   result of the training of Copilot or its suggestions, let alone any specific allegations supporting

12   such a theory. Their "[t]hreadbare recital[] of the elements of [the] cause of action, supported by

13   mere conclusory statements, do[es] not suffice." *Iqbal*, 556 U.S. at 678.

14   In all events, Plaintiffs' UCL claim fails because it is dependent upon otherwise defective

15   DMCA, tort, passing off, and privacy violations. "[T]o state a claim under the unlawful prong of

16   the UCL, a plaintiff must sufficiently plead a predicate violation." *Lopez v. Bank of Am., N.A.*,

17   505 F. Supp. 3d 961, 976 (N.D. Cal. 2020) (quotation marks omitted). Plaintiffs did not.

18   **V.   PLAINTIFFS' REVERSE PASSING OFF CLAIM IS BARRED UNDER *DASTAR*.**

19   Plaintiffs purport to assert a claim under the Lanham Act for "reverse passing off."

20   Compl. at 43 (Count V). The allegation that Defendants have marketed, distributed, or sold

21   works in which Plaintiffs claim some interest "without attribution" is a copyright claim

22   masquerading as a trademark claim, and therefore squarely barred by the Supreme Court's

23   decision in *Dastar Corp. v. Twentieth Century Fox Film Corp.*, 539 U.S. 23, 37 (2003); *see also*

24   *Sybersound Records, Inc. v. UAV Corp.*, 517 F.3d 1137, 1143-44 (9th Cir. 2008).

25   **VI.   PLAINTIFFS' PRIVACY CLAIMS AGAINST GITHUB FAIL.**

26   Plaintiffs attempt to claim that GitHub invaded their privacy (Counts VIII, IX, and X) at

27   the same time that the overall thrust of the Complaint is a failure to publicly attribute their

28   identity to code. The Complaint does not identify any specific privacy-related provision of any

1   alleged contract that was violated or how; does not allege a data breach that could yield liability

2   under the California Consumer Privacy Act ("CCPA"); and does not identify any personal data

3   that was negligently handled nor how that occurred. As noted in Part I.B, *supra*, no facts in the

4   Complaint explain how Plaintiffs were harmed by any privacy-based conduct by GitHub.

5   Whatever privacy-based claim Plaintiffs purport to advance, at a minimum they must

6   provide a description of the PII at issue and a plausible allegation of some conduct in connection

7   with that PII. Plaintiffs fail to provide those basics. *See* Compl. ¶¶ 215-39. The Complaint uses

8   the conclusory terms "personal data," "personal information," and "PII" in describing Plaintiffs'

9   claims, *e.g.*, Compl. ¶¶ 220, 228-29, but never defines those terms or specifies the actual data

10   types discussed. Similarly, the Complaint insinuates that GitHub is collecting, distributing, using,

11   or selling such information, but does so without alleging any facts. *See* Compl. ¶¶ 220, 229-33,

12   236-37. Plaintiffs never identify which personal information is involved, how it is captured,

13   shared, sold, or otherwise distributed, or how any of this has resulted or is likely to result in

14   concrete injury. Plaintiffs thus fail to provide basic notice of their claims for breach of GitHub's

15   Privacy Policy, violation of the CCPA, or negligence, necessitating dismissal.

16   The CCPA claim, moreover, faces a particularly demanding standard that Plaintiffs do not

17   even attempt to satisfy. The CCPA provides a limited private right of action. *See* Cal. Civ. Code

18   § 1798.150. It allows a private claim only if a consumer's "nonencrypted or nonredacted

19   personal information, as defined in [§ 1798.81.5(d)(1)(A)], is subject to an unauthorized access

20   and exfiltration, theft, or disclosure as a result of the business's violation of the duty to implement

21   and maintain reasonable security procedures and practices appropriate to the nature of the

22   information to protect the personal information." It is thus § 1798.81.5(d)(1)(A)—not

23   § 1798.140(o)(1), as Plaintiffs allege, Compl. ¶ 228—that defines personal information for

24   purposes of a CCPA claim. That provision defines such information as last name and first name

25   or first initial, *plus* one or more of a social security number, driver's license or similar number,

26   account number or card number with access code or password that would give access to a

27   financial account, medical information, health insurance information, biometric data, or genetic

28   data. *See* Cal. Civ. Code § 1798.81.5(d). Plaintiffs' allegations of misuse of unspecified PII are

GITHUB AND MICROSOFT'S MTD
NO. 4:22-CV-6823-JST

1 not actionable under § 1798.150, because Plaintiffs fail to plead any facts regarding data security

2 or theft of the specified categories of data that could support a CCPA claim. Simply parroting the

3 language of the statute is not enough to plausibly plead a claim, *see Cork v. CC-Palo Alto, Inc.*,

4 534 F. Supp. 3d 1156, 1183 (N.D. Cal. 2021), and that is particularly true here where the facts

5 described in the Complaint regarding training a machine learning model on publicly available

6 data appears to be far afield from the ordinary fact patterns in data breach cases.

7 **VII.   PLAINTIFFS' UNJUST ENRICHMENT, CONSPIRACY, AND DECLARATORY**
        RELIEF ALLEGATIONS FAIL ALONG WITH THE UNDERLYING CLAIMS.

8

9   Plaintiffs purport to advance an unjust enrichment claim against GitHub (Count VI), but

10 "there is no cause of action in California for unjust enrichment." *Melchior v. New Line Prods.,*

11 *Inc.*, 106 Cal. App. 4th 779, 793 (2003); *Bosinger v. Belden CDT, Inc.*, 358 F. App'x 812, 815

12 (9th Cir. 2009). Nor is civil conspiracy (Count XI) a cause of action under California law.

13 *Applied Equip. Corp. v. Litton Saudi Arabia Ltd.*, 7 Cal. 4th 503, 514 (1994).

14   Plaintiffs' declaratory relief claim (Count XII) rests on the viability of underlying claims,

15 and should be dismissed along with them. Count XII also runs afoul of Article III because the

16 Complaint does not set forth a specific requested declaration, let alone make allegations

17 demonstrating an injury that would be redressed by a declaration. *See* Part I, *supra*; *Lujan*, 504

18 U.S. at 561; *see Hobbs v. Sprague*, 87 F. Supp. 2d 1007, 1012 (N.D. Cal. 2000) (no standing

19 where Complaint fails to show that an alleged injury would be redressed by a favorable decision).

20   **CONCLUSION**

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

22
    Dated: January 26, 2023   Orrick, Herrington & Sutcliffe LLP
23

24

25   By: _____*/s/ Annette L. Hurst*_____
                                                ANNETTE L. HURST
26   Attorneys for Defendants
                                            GitHub, Inc. and Microsoft Corp.
27

28