

No. 24-6136

---

IN THE  
**United States Court of Appeals for the Ninth Circuit**

---

J. DOE, 1; J. DOE, 2-5,  
individually and on behalf of all others similarly situated,  
*Plaintiffs-Petitioners,*

*v.*

GITHUB, INC., a Delaware corporation; MICROSOFT CORPORATION, a  
Washington corporation; OPENAI, INC., a Delaware nonprofit  
corporation; OPENAI, LP, a Delaware limited partnership; OPENAI GP,  
LLC, a Delaware limited liability company; OPENAI STARTUP FUND GP I,  
LLC, a Delaware limited liability company; OPENAI STARTUP FUND I,  
LP, a Delaware limited partnership; OPENAI STARTUP FUND  
MANAGEMENT, LLC, a Delaware limited liability company; OPENAI  
OPCo, LLC; OPENAI, LLC; OPENAI GLOBAL, LLC; OAI CORPORATION;  
OPENAI HOLDINGS, LLC; OPENAI HOLDCO, LLC; OPENAI INVESTMENT;  
OPENAI STARTUP FUND SPV I, LP; OPENAI STARTUP FUND SPV GP I,  
LLC,  
*Defendants-Respondents.*

---

Petition for Permission to Appeal an Order from the  
United States District Court for the Northern District of California  
No. 4:22-cv-06823-JST, Hon. Jon S. Tigar

---

**RESPONSE TO PETITION FOR PERMISSION TO APPEAL  
ORDER UNDER 28 U.S.C. § 1292(b)  
(REDACTED)**

---

Christopher J. Cariello  
ORRICK, HERRINGTON &  
SUTCLIFFE LLP  
51 West 52nd Street  
New York, NY 10019  
(212) 506-5000

Annette L. Hurst  
ORRICK, HERRINGTON &  
SUTCLIFFE LLP  
405 Howard Street  
San Francisco, CA 94105  
(415) 773-5700

*Additional Counsel Listed on Inside Cover*

William W. Oxley  
Alyssa M. Caridis  
ORRICK, HERRINGTON &  
SUTCLIFFE LLP  
355 S. Grand Avenue, Suite 2700  
Los Angeles, CA 90071

*Counsel for Defendants-Respondents  
GitHub, Inc. and Microsoft Corporation*

## CORPORATE DISCLOSURE STATEMENT

Pursuant to Federal Rule of Appellate Procedure 26.1, GitHub, Inc. states that Microsoft Corporation is its parent company and owns 10% or more of GitHub's stock. Microsoft Corporation hereby states that it has no parent company, and no publicly-held corporation owns 10% or more of its stock.

ORRICK, HERRINGTON & SUTCLIFFE LLP

*/s/ Annette L. Hurst*

---

Annette L. Hurst

*Counsel for Defendants-Respondents*

*GitHub, Inc. and Microsoft Corporation*

## TABLE OF CONTENTS

	<b>Page</b>
CORPORATE DISCLOSURE STATEMENT.....	i
TABLE OF AUTHORITIES.....	iii
INTRODUCTION.....	1
STATEMENT OF THE CASE .....	5
GitHub Copilot Is A Software Code Completion Tool Based On Generative AI.....	5
Plaintiffs Challenge Copilot, But Most Of Their Claims Are Dismissed. ....	6
The District Court Dismisses Plaintiffs’ § 1202(b) Claims With Prejudice But Certifies Its Decision Under § 1292(b).....	7
REASONS FOR DENYING THE PETITION .....	9
An Interlocutory Appeal Is Not Justified By The § 1292(b) Factors And This Case Is Poorly Suited To Interlocutory Review. ....	9
A. There is no substantial ground for difference of opinion on a controlling issue of law. ....	10
1. An overwhelming consensus supports the rule that a § 1202(b) claim requires removal of CMI from a substantially identical copy. ....	10
2. There is not a substantial divergence among the decisions of other courts. ....	16
B. This case is a poor candidate for interlocutory review because of standing defects and alternative grounds for affirmance.....	19
1. Plaintiffs lack standing for their § 1202(b) claims. ....	19
2. Plaintiffs’ § 1202(b) claims fail on independent grounds.....	23
CONCLUSION.....	27
CERTIFICATE OF COMPLIANCE	

## TABLE OF AUTHORITIES

	Page(s)
<b>Cases</b>	
<i>ADR Int’l Ltd. v. Inst. for Supply Mgmt. Inc.</i> , 667 F. Supp. 3d 411 (S.D. Tex. 2023) .....	16
<i>Advanta-STAR Auto. Rsch. Corp. of Am. v. Search Optics, LLC</i> , 672 F. Supp. 3d 1035 (S.D. Cal. 2023).....	11
<i>Andersen v. Stability AI Ltd.</i> , No. 23-cv-00201, 2024 WL 3823234 (N.D. Cal. Aug. 12, 2024).....	11, 17
<i>Beijing Meishe Network Tech. Co. v. TikTok Inc.</i> , No. 23-cv-06012-SI, 2024 WL 3522196 (N.D. Cal. July 23, 2024) .....	17
<i>Bounce Exch., Inc. v. Zeus Enter., Ltd</i> , No. 15-cv-3268 (DLC), 2015 WL 8579023 (S.D.N.Y. Dec. 9, 2015).....	18
<i>Buchholz v. Meyer Njus Tanick, PA</i> , 946 F.3d 855 (6th Cir. 2020).....	21
<i>In re Cement Antitrust Litig. (MDL No. 296)</i> , 673 F.2d 1020 (9th Cir. 1981).....	9, 10
<i>Clapper v. Amnesty Int’l USA</i> , 568 U.S. 398 (2013).....	21
<i>Couch v. Telescope Inc.</i> , 611 F.3d 629 (9th Cir. 2010).....	9, 10, 17
<i>Design Basics, LLC v. WK Olson Architects, Inc.</i> , No. 17 C 7432, 2019 WL 527535 (N.D. Ill. Feb. 11, 2019) .....	12
<i>Falkner v. Gen. Motors LLC</i> , 393 F. Supp. 3d 927 (C.D. Cal. 2018) .....	12

<i>Faulkner Press, L.L.C. v. Class Notes, L.L.C.</i> , 756 F. Supp. 2d 1352 (N.D. Fla. 2010) .....	12
<i>Free Speech Sys., LLC v. Menzel</i> , 390 F. Supp. 3d 1162 (N.D. Cal. 2019) .....	23
<i>Frost-Tsuji Architects v. Highway Inn, Inc.</i> , No. 13-00496, 2015 WL 263556 (D. Haw. Jan. 21, 2015) .....	11
<i>Harrington v. Pinterest, Inc.</i> , No. 20-cv-05290, 2021 WL 4033031 (N.D. Cal. Sept. 3, 2021).....	25
<i>James v. Price Stern Sloan, Inc.</i> , 283 F.3d 1064 (9th Cir. 2002) .....	4, 9
<i>Kadrey v. Meta Platforms, Inc.</i> , No. 23-CV-03417-VC, 2023 WL 8039640 (N.D. Cal. Nov. 20, 2023).....	23
<i>Kelly v. Arriba Soft Corp.</i> , 77 F. Supp. 2d 1116 (C.D. Cal. 1999) .....	12
<i>Kirk Kara Corp. v. W. Stone &amp; Metal Corp.</i> , No. CV 20-1931, 2020 WL 5991503 (C.D. Cal. Aug. 14, 2020) .....	11
<i>Lujan v. Defs. of Wildlife</i> , 504 U.S. 555 (1992).....	22
<i>Nat. Res. Def. Council v. U.S. Env't Prot. Agency</i> , 857 F.3d 1030 (9th Cir. 2017) .....	12
<i>Oracle Int'l Corp. v. Rimini St., Inc.</i> No. 19-cv-01987, 2023 WL 4706127 (D. Nev. July 24, 2023).....	18
<i>Philpot v. Alternet Media, Inc.</i> , No. 18-cv-4479, 2018 WL 6267876 (N.D. Cal. Nov. 30, 2018) .....	25
<i>Pinkert v. Schwab Charitable Fund</i> , 48 F.4th 1051 (9th Cir. 2022) .....	20
<i>Real World Media LLC v. Daily Caller, Inc.</i> , No. 23-cv-1654, 2024 WL 3835351 (D.D.C. Aug. 14, 2024).....	17, 18

<i>Software Pricing Partners LLC v. Geisman</i> , No. 19-cv-00195, 2022 WL 3971292 (W.D.N.C. Aug. 31, 2022) .....	18
<i>Splunk, Inc v. Cribl, Inc.</i> , 662 F. Supp. 3d 1029 (N.D. Cal. 2023) .....	18
<i>Steel Co. v. Citizens for a Better Env't</i> , 523 U.S. 83 (1998).....	19
<i>Stevens v. Corelogic, Inc.</i> , 899 F.3d 666 (9th Cir. 2018).....	24
<i>Textile Secrets Int'l, Inc. v. Ya-Ya Brand Inc.</i> , 524 F. Supp. 2d 1184 (C.D. Cal. 2007) .....	12
<i>TransUnion LLC v. Ramirez</i> , 594 U.S. 413 (2021).....	21
<i>Tremblay v. OpenAI, Inc.</i> , 716 F. Supp. 3d 772 (N.D. Cal. 2024) .....	11, 25

## Statutes

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

§ 1201.....	14
§ 1201(d)(2).....	14
§ 1202.....	6, 12, 14, 15
§ 1202(b) .....	<i>passim</i>
§ 1202(c).....	11
28 U.S.C. § 1292(b).....	4, 7, 8, 9, 17

## INTRODUCTION

This petition for interlocutory review is an advisory opinion waiting to happen. Rushing to become the first class-action litigants attacking groundbreaking generative AI technology, Plaintiffs asserted a dozen causes of action—everything from breach of contract to privacy torts—against GitHub’s Copilot source-code-completion tool. Among the ten that have now been dismissed are claims under 17 U.S.C. § 1202(b), one of the Digital Millennium Copyright Act’s (“DMCA”) provisions protecting “Copyright Management Information,” or “CMI.” Plaintiffs hypothesize that Copilot may sometimes generate a short snippet of suggested software code similar to existing copyrighted code, without also reproducing authorship and license information that constitutes CMI. But conjecture is not a claim, and Plaintiffs rushed to Court with no more.

Plaintiffs’ first problem is that they cannot plausibly allege that the theoretical injury they posit has happened, or is likely to happen, to them. As their Complaint<sup>1</sup> details, Copilot is no copy-paste machine.

---

<sup>1</sup> We cite the operative Second Amended Complaint, an unredacted version of which is attached as Exhibit A to the provisionally sealed



When a Copilot user is coding, Copilot *generates* context-specific coding suggestions in real-time based on “inferred ... statistical patterns.”

SAC ¶ 64. Plaintiffs allege that a mere “1%” of suggestions will match *someone’s* preexisting code, Dkt. 1-1 at 181—itself a remote prospect and largely a function of the weight the model gives to the sort of boilerplate code “found in the most projects,” SAC ¶ 91. But even worse, Plaintiffs have never plausibly alleged any likelihood that a Copilot suggestion would match *their* code. This leaves them without Article III standing or a viable claim.

A second problem is that Plaintiffs have never even *identified* what the software code at issue in this case is. Section 1202(b) is an anti-piracy statute focused on “remov[al]” of CMI—“information conveyed *in connection with copies ... of a work.*” Yet Plaintiffs do not tell us what works or copies are at issue. Section 1202(b) also requires that a defendant have reason to know removal of CMI will “induce, enable, facilitate, or conceal an infringement.” Yet Plaintiffs point to no allegedly infringed code. Indeed, one of the great oddities of this case is

---

version of this opposition, as “SAC” and refer to it as the “Complaint.” We cite Plaintiffs’ petition as “Pet.” and the docket in this Court as “Dkt.”

that Plaintiffs have never even tried to advance a copyright infringement claim at all. Plaintiffs' failure to identify the works at issue not only undermines any § 1202(b) claim—it makes for a case where everyone is shooting in the dark.

The district court dismissed Plaintiffs § 1202(b) claims on a different ground, which is the one on which Plaintiffs now seek review. Even if Plaintiffs could allege that Copilot has emitted something like *their* code, and even if it were clear what that code *is*, Plaintiffs at most allege that Copilot outputs can be inexact matches for short snippets of code. But as a host of district courts have held, a § 1202(b) claim requires a demonstration that CMI has been removed from a *substantially identical* copy of a work. Otherwise, the plaintiff complains not about the active “remov[al]” of CMI from a “copy” of a work that § 1202(b) addresses, but of the mere failure to *reproduce* CMI alongside some *variation* on a work.

There is no good reason to review this decision at this stage and prudence counsels against it.

Courts have repeatedly endorsed the substantial identity rule on which the district court's decision rests, with nothing but scattered

outliers. An issue that has come up repeatedly and been resolved the same way almost every time is not “novel and difficult.” And while Plaintiffs pick at semantic differences across a few cases that have not demanded literal identicality, they point to *no* case approving a rule that would allow their nebulous claims to survive. *Infra* § A.

In any event, even if the issue needed clarifying, this case is a poor vehicle for doing so. To begin with, this Court *cannot* reach the issue because Plaintiffs lack Article III standing. Even if Plaintiffs overcame that hurdle, a ruling on the substantial-identicality rule would be academic, because multiple alternative grounds support dismissal of Plaintiffs’ dubious claims. Far from a clean opportunity to resolve the discrete legal issue Plaintiffs purport to frame, review here would demand the sort of full-blown appeal best addressed after final judgment. And far from clarifying the law for other AI cases, a pleading-stage precedent based on Plaintiffs’ under-articulated claims would generate far more heat than light. *Infra* § B.

This Court denies petitions to appeal under § 1292(b) “quite frequently,” *James v. Price Stern Sloan, Inc.*, 283 F.3d 1064, 1067-68 n.6 (9th Cir. 2002), and it should do so here.

## STATEMENT OF THE CASE

### *GitHub Copilot Is A Software Code Completion Tool Based On Generative AI.*

Over the past several years, scientists and engineers have made dramatic advances in generative artificial intelligence. GitHub Copilot, a generative AI tool offered by the popular open source and software development platform GitHub, is built on an AI model created by Defendant OpenAI. Such models are typically trained through exposure to a corpus of material called “training data.” SAC ¶ 93.

To create the model underlying Copilot, OpenAI included in the training data software code that GitHub users like Plaintiffs make publicly available on GitHub. SAC ¶ 181. Through exposure to large amounts of code, the model discerns statistical patterns that become part of the model. SAC ¶ 93. When built into GitHub Copilot, the model is then capable of “predic[ting] ... the most likely [coding] solution,” “suggest[ing] code and entire functions in real-time” for incorporation into a coding project. SAC ¶ 59, 91.

Copilot thus leverages software code that developers have shared with the public, enabling those same developers to build upon that collective learning to create new projects. It advances the profound

principles of “open source software”—a concept that encourages developers to use, study, change, and share source code for the public good.

***Plaintiffs Challenge Copilot, But Most Of Their Claims Are Dismissed.***

In November 2022, Plaintiffs (proceeding pseudonymously), filed a putative class-action against GitHub, Microsoft (which owns GitHub), and OpenAI. They asserted causes of action for breach of various open-source licenses and GitHub user agreements, violations of § 1202, and claims under other federal and state laws. They did not allege copyright infringement.

The first two rounds of motions to dismiss narrowed Plaintiffs’ case considerably. The district court found that Plaintiffs had not alleged an injury-in-fact to establish Article III standing for privacy claims. Dkt. 1-1 at 179. The district court also found that Plaintiffs’ claims under the Lanham Act and various state-law theories failed. Dkt. 1-1 at 194-97; Dkt. 1-1 at 280-85.

This left two sets of claims: (a) breach-of-contract claims that remain pending before the district court and are not at issue here; and (b) claims under § 1202(b) asserting that Copilot may emit code

suggestions that are inexact matches of small portions of code on which the model underlying Copilot was trained.

***The District Court Dismisses Plaintiffs' § 1202(b) Claims With Prejudice But Certifies Its Decision Under § 1292(b).***

Defendants sought dismissal of the § 1202(b) claims on multiple grounds. To begin with, Plaintiffs had failed to establish “that they themselves ha[d] suffered the injury they describe[d].” Dkt. 1-1 at 180. Initially, the district court agreed, finding that Plaintiffs lacked standing to pursue damages. Dkt. 1-1 at 180. But it held that there is a “substantial risk” that Copilot “will reproduce Plaintiffs’ licensed code as output” based solely on allegations that Plaintiffs’ code was in the training set and that “about 1% of the time” output matches something “from training data.” Dkt. 1-1 at 181.

In their First Amended Complaint, Plaintiffs tried a new tactic to establish damages standing: Self-inflicted injury. For three of the Doe Plaintiffs, they offered “examples” where they were allegedly able to enter lines of their own code into a GitHub coding project, and this “prompt[ed]” Copilot to emit short code “suggest[ions]” that resembled the next several lines of the Does’ code. N.D. Cal. ECF No. 97-3 (FAC) ¶¶ 98, 100-03. Plaintiffs did not allege facts suggesting that any real

user has ever, could ever, or would ever want to enter these prompts.

But the district court nevertheless found that the manufactured injury sufficed to establish standing to pursue damages. Dkt. 1-1 at 279.

Plaintiffs could pursue damages, however, only as to their breach-of-contract claims, because the district court dismissed the § 1202(b) claims. Pointing to numerous cases holding that “no DMCA violation exists where the works are not identical,” Dkt. 1-1 at 286 (internal quotation marks and citation omitted), the court found that Plaintiffs had at most alleged Copilot-generated outputs that were “modified,” a “variation[],” or the “functional[] equivalent” of Plaintiffs’ unidentified works, *id.* (quoting FAC ¶¶ 103, 110, 120). After Plaintiffs’ Second Amended Complaint failed to fix the defect, the district court dismissed the § 1202(b) claims with prejudice. Dkt. 1-1 at 33 (finding that Plaintiffs could not “identify even a single example of Copilot producing an identical copy of any work”).

Though Plaintiffs’ core breach-of-contract claims remained in the case, they moved the district court to certify its order for interlocutory appeal under § 1292(b) and to stay discovery. Dkt. 1-1 at 290-310. Plaintiffs were candid about their reason: The § 1292(b) claim

represented “the bulk of [their] damages,” Dkt. 1-1 at 298. The district court granted the motion in a brief order. Dkt. 1-1 at 45-49. It pointed to three district court cases purportedly declining to apply a substantial-identity standard to DMCA claims. Dkt. 1-1 at 47. And it found that resolving the issue “would materially advance ... this particular case” and also “others in the Circuit raising the same issue.” *Id.* The district court stayed the litigation pending any appeal.

### **REASONS FOR DENYING THE PETITION**

#### **An Interlocutory Appeal Is Not Justified By The § 1292(b) Factors And This Case Is Poorly Suited To Interlocutory Review.**

“[S]ection 1292(b) is to be applied sparingly and only in exceptional cases.” *In re Cement Antitrust Litig. (MDL No. 296)*, 673 F.2d 1020, 1027 (9th Cir. 1981). Because § 1292(b) is a “departure from the normal rule that only final judgments are appealable,” it “must be construed narrowly.” *James v. Price Stern Sloan, Inc.*, 283 F.3d 1064, 1067 n.6 (9th Cir. 2002). “[T]he party pursuing the interlocutory appeal bears the burden” of establishing each jurisdictional element. *Couch v. Telescope Inc.*, 611 F.3d 629, 633 (9th Cir. 2010). And even where a district court certifies its order, this Court evaluates § 1292(b)’s



jurisdictional factors to ensure they are met, and in all events retains discretion to deny review. *In re Cement*, 673 F.2d at 1026.

Here, the statutory factors are not satisfied and multiple prudential considerations counsel against interlocutory review. There is not a substantial or mature disagreement among other courts on the requirement of substantial identity. And the district court's order will not provide a clean opportunity for the Court to consider this question because of standing defects and alternative grounds for dismissal.

**A. There is no substantial ground for difference of opinion on a controlling issue of law.**

**1. An overwhelming consensus supports the rule that a § 1202(b) claim requires removal of CMI from a substantially identical copy.**

Plaintiffs assert that whether § 1202(b) claims require removal of CMI from a substantially identical copy of a work is a question with a “substantial ground for difference of opinion.” Pet. 7. But they do not (and cannot) claim that “the circuits are in dispute” on this point, and they come nowhere close to showing that it presents a “novel and difficult question[] of first impression.” *Couch*, 611 F.3d at 633 (citation omitted); see Pet. 7-8. Just the opposite, courts have routinely rejected

§ 1202(b) claims that, like Plaintiffs', challenge a defendant's mere *non-inclusion* of CMI along with some modified, excerpted, or adapted version of the original.

What is sometimes dubbed the “identity” requirement in fact reflects a broader throng of decisions reinforcing § 1202(b)'s requirement of “remov[al]” of CMI, which is defined as “information conveyed in connection with *copies ... of a work*,” § 1202(c) (emphasis added). Many courts have invoked a requirement of “identical” copies in rejecting claims based on purported removal of CMI from inexact imitations, tracings, and more recently the short, generated output of generative AI tools.<sup>2</sup> A related set of cases, while not phrasing their rulings in terms of identical copies, rejects CMI-removal claims based on mere excerpts, thumbnails, or framings of works, finding that only

---

<sup>2</sup> *E.g.*, *Kirk Kara Corp. v. W. Stone & Metal Corp.*, No. CV 20-1931, 2020 WL 5991503, at \*6 (C.D. Cal. Aug. 14, 2020); *Advanta-STAR Auto. Rsch. Corp. of Am. v. Search Optics, LLC*, 672 F. Supp. 3d 1035, 1057 (S.D. Cal. 2023); *Frost-Tsuji Architects v. Highway Inn, Inc.*, No. 13-00496, 2015 WL 263556, at \*3 (D. Haw. Jan. 21, 2015), *aff'd*, 700 F. App'x 674 (9th Cir. 2017); *Tremblay v. OpenAI, Inc.*, 716 F. Supp. 3d 772, 779-80 (N.D. Cal. 2024) (generative AI); *Andersen v. Stability AI Ltd.*, No. 23-cv-00201, 2024 WL 3823234, at \*8 (N.D. Cal. Aug. 12, 2024) (generative AI).

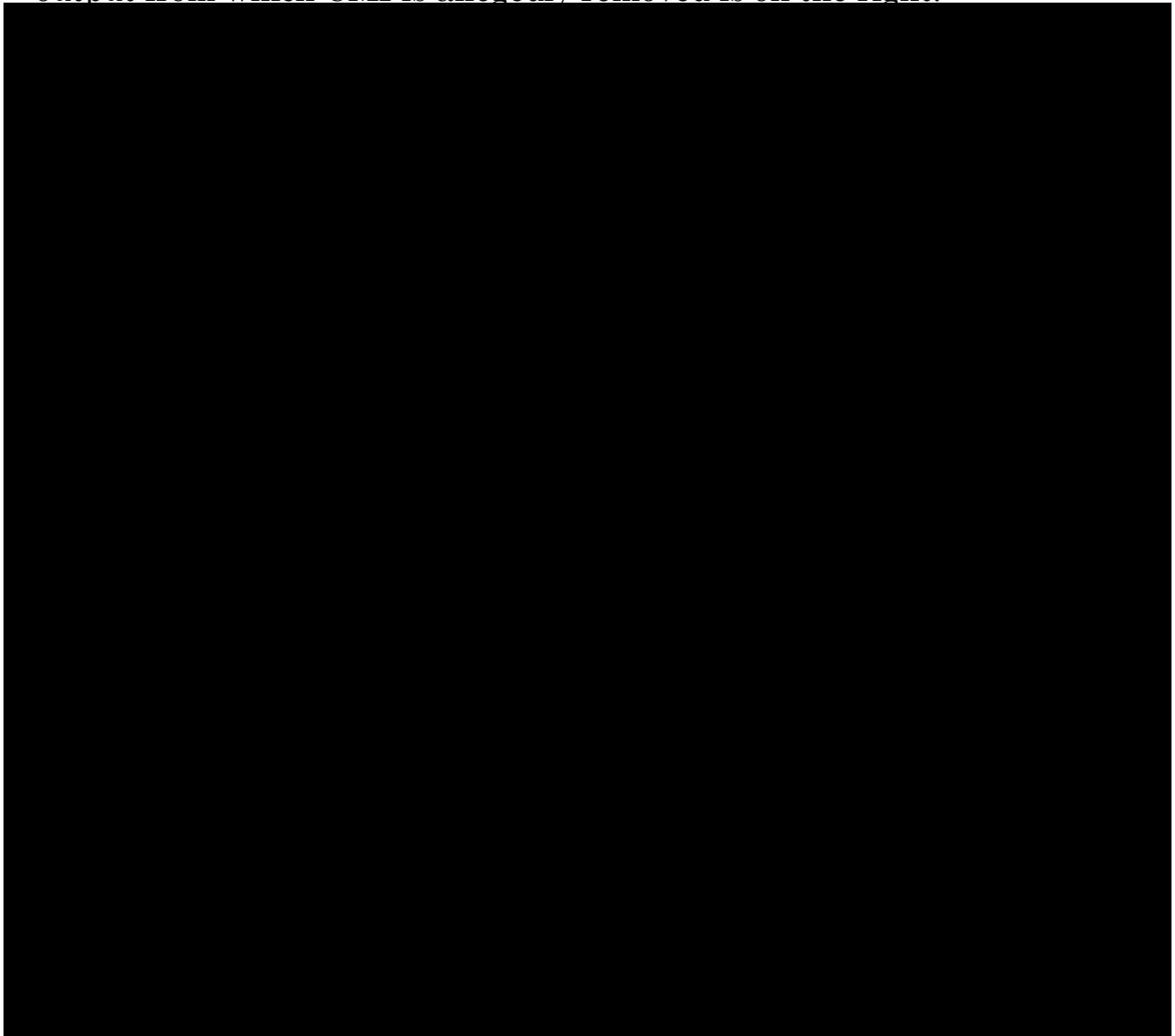
removal of CMI from an actual copy of a work suffices.<sup>3</sup>

All of these rulings reflect a natural reading of § 1202(b). Again, the provision protects the integrity of information conveyed “with a copy.” It makes no sense to speak of *removing* CMI from a “variation[]” or “functional[] equivalent,” Dkt. 1-1 at 286-87 (Order on Motions to Dismiss FAC)—CMI was never included with those things in the first place. And the reading makes all the more sense “in light of the ... language and design of the statute as a whole,” *Nat. Res. Def. Council v. U.S. Env’t Prot. Agency*, 857 F.3d 1030, 1036 (9th Cir. 2017). Section 1202 is one of two adjoining provisions from the DMCA concerned with the “rapid dissemination of perfect copies,” *Textile Secrets Int’l, Inc. v. Ya-Ya Brand Inc.*, 524 F. Supp. 2d 1184, 1198 (C.D. Cal. 2007) (quoting House Committee Report). It is not a sweeping new attribution right requiring the reproduction of CMI any time some new content bears some resemblance to an original.

---

<sup>3</sup> *Falkner v. Gen. Motors LLC*, 393 F. Supp. 3d 927, 938-39 (C.D. Cal. 2018); *Faulkner Press, L.L.C. v. Class Notes, L.L.C.*, 756 F. Supp. 2d 1352, 1356, 1359 (N.D. Fla. 2010) (excerpts); *Design Basics, LLC v. WK Olson Architects, Inc.*, No. 17 C 7432, 2019 WL 527535, at \*5 (N.D. Ill. Feb. 11, 2019); *Kelly v. Arriba Soft Corp.*, 77 F. Supp. 2d 1116, 1122 (C.D. Cal. 1999), *rev’d on other grounds*, 336 F.3d 811 (9th Cir. 2003).

Plaintiffs' claims were correctly dismissed under the above consensus. Consider the following example that Plaintiffs coaxed out of Copilot by entering verbatim *twenty-two lines* of their own code as a prompt—code you could only enter if you already had the full work to begin with. Some of Doe 1's actual code is on the left, and the Copilot output from which CMI is allegedly removed is on the right:



SAC ¶¶ 118-19.

A glance reveals plain differences in length, structure, and syntax, and careful scrutiny reveals still more variation. The newly-generated code on the right is, at most, functionally similar to that on the left. There is no reading of § 1202(b) under which the output on the right, generated on the fly by Copilot based on user prompting, could be considered a “copy” of the passage on the left—let alone a copy of the full “work” in connection with which CMI was originally included. A claim based on Doe 1’s example would thus fail under any of the cases cited above, whether those adopting an explicit substantial identity rule or those otherwise rejecting claims based on excerpts, modifications, and the like.

Plaintiffs resist this by seizing on the word “identical,” noting that it does not appear in § 1202(b). Pet. 17. They also note that a neighboring provision addressing entirely different subject matter—§ 1201, pertaining to “[c]ircumvention” of “technological measures”—uses the phrase “identical copy” to define an obscure exemption for “nonprofit libraries, archives, and educational institutions.” 17 U.S.C. § 1201(d)(2). From this, they argue that “if Congress had intended to require ‘identity’ for § 1202 liability, it would have said so.” Pet. 17.

But Plaintiffs’ narrow focus on the word “identical” is largely a distraction here—no literal identity requirement is needed to dispose of claims like the example above. As the Complaint details, Copilot does not copy; it *generates* short code suggestions based on its understanding of inferred statistical relationships across an immense body of publicly available code. *See* SAC ¶¶ 8, 58, 60-66, 80-85. Even if these snippets are “functional equivalents” of existing code, Copilot does not output copies of any complete “work,” and therefore cannot be removing *anything* from *any* copy of such a “work.” So even if this Court accepted review and agreed that precise “identity” is not required, Plaintiffs’ claims fail just as easily under the line of cases that do not use that word, but reject claims attacking excerpts, reframings, and the like. *Supra* n.3.

For similar reasons, Plaintiffs’ textual arguments fail. Though Congress did not use the word “identical” in § 1202, it also did not use words like “alteration,” “modification,” “abridgement,” “adaptation,” “derivative,” or any number of others that would expand § 1202(b) to reach the sorts of functionally similar content Plaintiffs allege. Its

decision not to do so is why court after court has rejected claims like Plaintiffs’.

**2. There is not a substantial divergence among the decisions of other courts.**

Plaintiffs claim that “[d]istrict courts in this Circuit and across the country are ... sharply divided” on whether § 1202(b) “requires that CMI be removed or altered from an identical copy.” Pet. 7-10. They overstate the disagreement. One out-of-circuit district court decision arguably conflicts with the district court’s decision here. The rest of the cases Plaintiffs cite reflect semantic disagreement over the term “identical.” And most importantly, *none* of these cases approve a § 1202(b) claim for the sorts of variations or functional equivalents on which Plaintiffs’ predicate their Complaint.

The one out-of-circuit district court decision that conflicts with the consensus is *ADR International Ltd. v. Institute for Supply Management Inc.*, 667 F. Supp. 3d 411 (S.D. Tex. 2023). Pet. 9. That case involved an unquestionable effort to create a copy of an entire work with only superficial modifications as “a thin veil over the facially obvious copying.” 667 F. Supp. 3d at 418-19. So it is not clear why the *ADR* court felt a need to consider the identity requirement at all.

Nor is it clear that the *ADR* court would use the same reasoning to approve Plaintiffs' claims arising from very different circumstances. But even if *ADR*'s unnecessary discussion were not dicta, it is a single out-of-circuit case arrayed against the weight of authority. A § 1292(b) appeal is not warranted simply because "counsel contends that one precedent rather than another is controlling." *Couch*, 611 F.3d at 633.

Apart from *ADR*, Plaintiffs trumpet three decisions published in the last several months, Pet. 10-12, but none is contrary to the district court's order here. The first, *Beijing Meishe Network Technology Co. v. TikTok Inc.*, did not hold that identity is not an element of liability under § 1202(b). It simply held that because of "factual differences between this case and the cases on which defendants rely, the Court cannot conclude as a matter of law at this pleading stage that plaintiff's claims should be dismissed on this basis." No. 23-cv-06012-SI, 2024 WL 3522196, at \*9 (N.D. Cal. July 23, 2024).

The second, *Andersen v. Stability AI Ltd.*, as noted (at *supra* n.2), expressly agrees with the district court's ruling here. The court noted that *ADR* held otherwise, but highlighted the unanimity "within the Ninth Circuit." 2024 WL 3823234, at \*8. And *Real World Media LLC*



*v. Daily Caller, Inc.*, expressly stated that it was not addressing the identical copies issue, calling it “a red herring here.” No. 23-cv-1654, 2024 WL 3835351, at \*10 (D.D.C. Aug. 14, 2024).<sup>4</sup>

So it is here, too. The vast majority of courts have adopted a substantial identity requirement. But even if some have not, Plaintiffs have pointed to no case that has ever endorsed a CMI-removal claim for anything like the short, modified code snippets Plaintiffs purport to challenge here. There is therefore no substantial ground for disagreement on a controlling issue of law.

---

<sup>4</sup> Plaintiffs invoke a handful of even less relevant cases involving allegations that a defendant directly copied copyrighted works and actively removed CMI, in which the identity question was not even squarely at issue. *See* Pet. 9-10. These cases do not move the needle. *See Oracle Int’l Corp. v. Rimini St., Inc.* No. 19-cv-01987, 2023 WL 4706127, at \*82 (D. Nev. July 24, 2023) (deliberate and direct copying; identical copies argument “improperly raised”); *Software Pricing Partners LLC v. Geisman*, No. 19-cv-00195, 2022 WL 3971292, at \*1-2 (W.D.N.C. Aug. 31, 2022) (default judgment in case alleging direct copying and alteration of proprietary documents; identical copies issue not presented); *Bounce Exch., Inc. v. Zeus Enter., Ltd*, No. 15-cv-3268 (DLC), 2015 WL 8579023, at \*2-4 (S.D.N.Y. Dec. 9, 2015); (copying and active removal of CMI from source code; identity issue not raised or discussed); *Splunk, Inc v. Cribl, Inc.*, 662 F. Supp. 3d 1029, 1053 (N.D. Cal. 2023) (same).

**B. This case is a poor candidate for interlocutory review because of standing defects and alternative grounds for affirmance.**

Even if the substantial-identity ruling were worthy of review, this case is a bad vehicle for doing so. Plaintiffs lack standing for their § 1202(b) claims, an issue that would prevent resolution on the merits. *See Steel Co. v. Citizens for a Better Env't*, 523 U.S. 83, 93-94 (1998) (standing must be addressed before the merits). Their § 1202(b) claims, moreover, fail on fully independent grounds, which would render any decision on the substantial-identity rule academic. And while Plaintiffs urge this Court to offer “guidance” as “AI-related cases continue to proliferate,” Pet. 15-16, it is hard to imagine more improvident circumstances than a pleading-stage interlocutory appeal by Plaintiffs without standing or a clearly articulated claim.

**1. Plaintiffs lack standing for their § 1202(b) claims.**

Plaintiffs’ § 1202(b) claims present a basic standing problem. Plaintiffs allege that their publicly available software code was some infinitesimal part of a vast training set used to train the model underlying Copilot—that is, their code was used as *input*. SAC ¶¶ 58, 109. But for purposes of their § 1202(b) claims, they assert no *input*-

based injury or claim. Their theory is that Copilot could *output* a suggestion that matches their code, without reproducing CMI. SAC ¶¶ 109, 112-138. The problem: They have never plausibly alleged that Copilot *ever has or would* output something resembling *their* code—i.e., that the injury has “materialized”—or that it is likely to output *their* code—i.e., that the injury is “certainly impending.” *Pinkert v. Schwab Charitable Fund*, 48 F.4th 1051, 1055 (9th Cir. 2022) (cleaned up). They just allege that this injury could hypothetically befall someone.

The district court committed legal error in concluding otherwise. First, in finding that three of the Doe Plaintiffs have standing to pursue damages, the Court relied on “examples” where Plaintiffs typed long sequences of their *own code* into Copilot, causing Copilot to produce output that is “most likely” to follow, SAC¶ 91—which, unsurprisingly, bears some functional resemblance to the way Plaintiffs’ ensuing code operates. Dkt. No. 1-1 at 278-79; *see* SAC ¶¶ 125-32. Plaintiffs do not allege that this has or could ever happen in the real world to users who do not already have Plaintiffs’ code and are not deliberately attempting to manufacture standing for a litigation. But the district court

nevertheless relied on Plaintiffs' self-inflicted injury as a basis for standing.

This is incorrect. Plaintiffs cannot “manufacture standing merely by inflicting harm on themselves.” *Clapper v. Amnesty Int’l USA*, 568 U.S. 398, 416 (2013). That is because “[a] self-inflicted injury, by definition, is not traceable to anyone but the plaintiff.” *Buchholz v. Meyer Njus Tanick, PA*, 946 F.3d 855, 866-67 (6th Cir. 2020) (collecting authorities). Were it otherwise, Article III’s injury requirement would be a nullity. For example, the Supreme Court held in *TransUnion LLC v. Ramirez* that the plaintiffs were not injured by the mere existence of inaccurate information in their credit files. 594 U.S. 413, 434 (2021). Under the district court’s approach, these plaintiffs could simply request their own credit report and that self-injury would confer standing.

Second, the district court found that Plaintiffs could seek injunctive relief based on their allegation that “GitHub has admitted that about 1% of the time, a suggestion may contain code snippets longer than ~150 characters that matches code from the training data.” SAC ¶ 207; *see also* SAC ¶¶ 102-03. But this could never establish

standing *as to these Plaintiffs*. All the allegation suggests is that 1% of Copilot’s outputs will contain a snippet matching *something* in the program’s vast training set—again, likely what the model has seen “in the most projects.” SAC ¶ 91. It most certainly does not mean that there is a 1% likelihood of a match with *everything* in the training set, nor that there is a 1% likelihood of a match with *Plaintiffs’* works. Plaintiffs’ thus fail to establish that they are “among the injured.” *Lujan v. Defs. of Wildlife*, 504 U.S. 555, 563 (1992).

Plaintiffs have had four Complaints to try to address this defect. To identify what their “works” are. To explain why those works might supply the sort of coding solution Copilot would suggest to a user. To articulate some plausible, real-world prompt that might yield an output replicating their code. Yet the most they have ever alleged is that their unidentified works are among the inputs for the model underlying Copilot, and that very occasionally an output matches *some* input. That is just the sort of “conjectural or hypothetical” injury that cannot confer standing. *Lujan*, 504 U.S. at 560. Granting review in this case would require this Court to confront the obvious standing problems barring Plaintiffs’ claims, preventing it from reaching the merits.

**2. Plaintiffs' § 1202(b) claims fail on independent grounds.**

If the Court did reach the merits, a meaningful decision would require evaluation of independent grounds for dismissal that would render any ruling on the substantial-identity rule academic. This Court should not grant review to entertain a full-blown appeal that will not change the result.

***No identified works.*** First, Plaintiffs fail to identify the copyrighted works at issue in their CMI claims. Courts have recognized such identification is required, because without it a court cannot evaluate whether plaintiffs have plausibly alleged removal of CMI from a copy of a work. *Free Speech Sys., LLC v. Menzel*, 390 F. Supp. 3d 1162, 1175 (N.D. Cal. 2019); *see also Kadrey v. Meta Platforms, Inc.*, No. 23-CV-03417-VC, 2023 WL 8039640, at \*2 (N.D. Cal. Nov. 20, 2023). The absence of identified works also makes it impossible for this Court to issue a substantive decision; review would be an exercise in abstraction.

Such an abstract ruling is particularly ill-advised here, where the Complaint casts serious doubt on whether Plaintiffs could ever base a claim on a copyrighted work. Consider the example Plaintiffs supplied

to Doe 2 in an attempt to generate standing. It involves “code” that is nothing but a list of the [REDACTED] SAC ¶¶ 112-16. A list of the [REDACTED] is hardly copyrightable material. And Plaintiffs’ complete failure to identify any copyrighted work strongly suggests that Plaintiffs’ claims invite a purely advisory opinion.

***Stevens v. Corelogic.*** Review of Plaintiffs’ § 1202(b) claims would also be incomplete without application of this Court’s decision in *Stevens v. Corelogic, Inc.*, 899 F.3d 666 (9th Cir. 2018). *Stevens* interpreted § 1202(b)’s requirement that a defendant remove CMI “knowing, or ... having reasonable grounds to know that it will induce, enable, facilitate, or conceal an infringement.” *Id.* at 673. Applying that requirement to technology that purportedly stripped CMI from works, this Court found that a plaintiff must show that “future infringement is *likely* ... to occur as a result of the removal or alteration of CMI.” *Id.* at 675 (emphasis added). The “general possibility” that CMI removal will lead to infringement is insufficient. *Id.* at 673.

Beyond threadbare allegations (e.g., SAC ¶¶ 221, 223), Plaintiffs have utterly failed to explain how infringement is likely to occur in the

context of their DMCA claims. This is unsurprising since Plaintiffs have never advanced a copyright infringement claim; do not identify any of their own copyrighted works; and offer no reason those works would ever be matched by Copilot output. *Stevens* thus forecloses Plaintiffs' claims.

The district court addressed the *Stevens* rule only briefly on Defendants' motion to dismiss the initial complaint, finding that it does not apply "at the pleading stage," Dkt. 1-1 at 192 (internal quotation marks omitted), before ultimately dismissing the § 1202(b) claims on identity grounds in later rounds of motion practice. The district court was wrong that *Stevens*'s objective inquiry cannot be applied at the pleading stage. See *Harrington v. Pinterest, Inc.*, No. 20-cv-05290, 2021 WL 4033031, at \*6 (N.D. Cal. Sept. 3, 2021); *Philpot v. Alternet Media, Inc.*, No. 18-cv-4479, 2018 WL 6267876, at \*5 (N.D. Cal. Nov. 30, 2018); *Tremblay*, 716 F. Supp. 3d at 779 (applying *Stevens* at pleading stage to generative AI tool). And in any event, when GitHub and Microsoft renewed the argument in connection with the Second Amended Complaint, the district court did not reach the issue.



If this Court were to grant review, full consideration of Plaintiffs' claims would require consideration of whether Plaintiffs satisfy the rule in *Stevens*. Otherwise the piecemeal appellate approach Plaintiffs seek would render a decision that itself delivers only a partial answer. The wiser approach is to decline interlocutory review that would sprawl well beyond the question the district court thought sufficient to justify an early appeal.

\*\*\*

Plaintiffs' final enticement is the promise that an early appellate decision here "would impact ... other cases," Pet. 15; *see* Pet. 12. That is a dangerous proposition in a case with claims as ill-defined as Plaintiffs'. Plaintiffs rushed tenuous § 1202(b) claims into court without standing to do so. They have spent the better part of two years amending their Complaint to plug holes in their § 1202(b) claims, yet still have never even identified the purported copyrighted works at issue—and in the software context, no less, where copyright protection is thinnest. So at best, review here would provide pleading-stage guidance for vaguely articulated § 1202(b) claims implicating unidentified software code. At worst, it would yield an advisory

precedent built on the shakiest of foundations. Because interlocutory appeal will not aid the sound resolution of this litigation or the development of the law, the Court should deny Plaintiffs' petition.

### CONCLUSION

The Court should deny Plaintiffs' petition to appeal.

October 17, 2024

Respectfully submitted,

William W. Oxley  
Alyssa M. Caridis  
ORRICK, HERRINGTON &  
SUTCLIFFE LLP  
355 S. Grand Avenue, Suite 2700  
Los Angeles, CA 90071

/s/ Annette L. Hurst  
Annette L. Hurst  
ORRICK, HERRINGTON &  
SUTCLIFFE LLP  
405 Howard Street  
San Francisco, CA 94105  
(415) 773-5700

Christopher J. Cariello  
ORRICK, HERRINGTON &  
SUTCLIFFE LLP  
51 West 52nd Street  
New York, NY 10019  
(212) 506-5000

*Counsel for Defendants-Respondents  
GitHub, Inc. and Microsoft Corporation*

## CERTIFICATE OF COMPLIANCE

The response complies with the type-volume limitation of Ninth Cir. R. 5-2(b) & 32-3(2) because this response contains 5195 words.

This response complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type style requirements of Fed. R. App. P. 32(a)(6) because this response has been prepared in a proportionally spaced typeface using Microsoft Word for Microsoft 365 in Century Schoolbook 14-point font.

ORRICK, HERRINGTON & SUTCLIFFE LLP

*/s/ Annette L. Hurst*

\_\_\_\_\_  
Annette L. Hurst

*Counsel for Defendants-Respondents  
GitHub, Inc. and Microsoft Corporation*