Ninth Circuit Rejects DMCA Section 1202(b) Theory Against Copilot Output in Doe v. GitHub

By Gregory Huh | September 18, 2026

RELATED PRACTICES Appellate and Supreme Court, Intellectual Property Litigation, AI and Technology, Copyright, Intellectual Property

AI Legal News

The Ninth Circuit affirmed on September 16, 2026 the dismissal of a Digital Millennium Copyright Act (DMCA) claim in *Doe v. GitHub, Inc.*, holding that programmers cannot sue over GitHub Copilot output that omits their attribution notices.[1] Judge Eric D. Miller wrote for a panel that also included Judge Sidney R. Thomas and District Judge Stanley Blumenfeld, Jr., sitting by designation. Section 1202(b) of the DMCA, the panel held, reaches affirmative acts against copyright management information attached to a work that already exists, while the complaint described a tool that creates "new works that never contained that information."

Background

The plaintiffs are anonymous programmers who published copyrighted, licensed code in public GitHub repositories. They sued GitHub, Microsoft and the OpenAI entities in the Northern District of California, alleging that Copilot and Codex — large language models trained on millions of GitHub projects — sometimes emit verbatim or near-verbatim copies stripped of the attribution and license terms that accompanied them. Two rounds of dismissal and amendment narrowed the case to one DMCA claim and two contract claims. Judge Jon S. Tigar dismissed the Section 1202(b) claim with prejudice, reasoning that the provision requires identical copies, and certified that question for interlocutory appeal under 28 U.S.C. Section 1292(b). The contract claims remain pending.

Analysis

Standing survives; the claim does not

The panel reached standing even though the certified order did not address it, reasoning that Section 1292(b) jurisdiction covers any issue fairly included within that order. It found the pleading sufficient: the complaint alleged that Copilot reproduces the named plaintiffs' code without attribution, cited research on models emitting memorized training data verbatim, and pointed to GitHub's own duplicate-detection filter for matching snippets of 150 characters or more as evidence that literal copying occurs. That made a substantial risk of injury plausible. The panel reserved whether the same material would carry the plaintiffs past summary judgment.

Identicality was the wrong label

More consequentially, the panel discarded the district court's vocabulary. Identicality is a misnomer it said: nothing in the statute demands literal sameness, and the question is what was done to the information. Section 1202(b) bars intentionally removing or altering that information, and both verbs presuppose material already attached to an existing work.[2] Identicality matters evidentially: where an accused copy matches an original in every respect but the missing notice, a factfinder may infer removal, as in *Friedman v. Live Nation Merchandise*.[3] Material differences point toward a new work the plaintiff's

© 2026 Haynes and Boone, LLP

information never accompanied. Minor cosmetic edits, the panel added, will not shield a defendant who substantially reproduces a work and deletes its notice.

A generative model is not a search engine

The panel then relied on the plaintiffs' own account of the technology. A system that infers statistical patterns and predicts the most likely completion is generating rather than retrieving, unlike a search engine returning stored copies of existing material. Retrieval-style output identical to the plaintiffs' code but missing the notice, the panel indicated, would have made a stronger claim. The court separately declined to consider an alternative "input" theory — that information was stripped during training — as forfeited, because counsel never corrected the district court's statement that the case was not about training. Treating substantial similarity as a Section 1202(b) violation, the panel warned, would let the DMCA supplant ordinary infringement law and its per-work damages cap.

Takeaways

- The output-side attribution theory is now foreclosed in the Ninth Circuit where most generative-AI copyright litigation sits, removing the DMCA's per-violation statutory damages as leverage there.
- The ruling narrows a claim, not a technology: the panel took no position on whether Copilot output infringes, and the open-source license contract claims continue in the district court.
- Standing was the plaintiffs' one win, resting on memorization research and GitHub's own duplicate filter — a pleading template other AI plaintiffs can reuse even where the merits fail.
- By recasting identicality as evidence rather than an element, the panel lines up with Judge Willett's August opinion for the Fifth Circuit, making a circuit split less likely.[4]

Endnotes

1. *Doe v. GitHub, Inc.*, No. 24-7700 (9th Cir. Sept. 16, 2026), available at https://cdn.ca9.uscourts.gov/datastore/opinions/2026/09/16/24-7700.pdf.
2. 17 U.S.C. Sections 1202 and 1203 (Copyright Act, Chapter 12), available at https://www.copyright.gov/title17/92chap12.html.
3. *Friedman v. Live Nation Merchandise, Inc.*, No. 14-55302 (9th Cir. Feb. 18, 2016), available at https://storage.courtlistener.com/pdf/2016/08/18/glen_friedman_v._live_nation_merchandise.pdf.
4. *Kipp Flores Architects, L.L.C. v. AMH Creekside Development, L.L.C.*, No. 23-50750 (5th Cir. Aug. 21, 2026) (opinion of Willett, J.), available at https://www.ca5.uscourts.gov/opinions/pub/23/23-50750-CV0.pdf.

© 2026 Haynes and Boone, LLP