Beyond Infringement: Rethinking DMCA § 1202 for Generative AI
78 STAN. L. REV. (forthcoming 2026)

NOTE

Beyond Infringement: Rethinking DMCA § 1202 for Generative AI

LARISSA BERSH[†]

Abstract. With the emergence of large language models (LLMs) like ChatGPT, scholars and courts have fervently debated whether LLMs' training on and reproduction of copyrighted materials amounts to fair use. But in a recent series of cases, a lesser-known challenge to LLMs has reared its head: § 1202 of the Digital Millenium Copyright Act. This provision requires that when a work is copied, its associated copyright management information (CMI)—such as its license or terms of use agreement—is copied with it. The requirement was originally intended to modernize copyright for the Internet by ensuring all users would be aware of the terms of their use. Now, § 1202's unintended overbreadth threatens to block LLM development and use as it swallows questions of infringement and fair use entirely.

This Note argues that § 1202 is broader than traditional copyright infringement in three critical respects: It allows liability without any showing of copyrightability, provides no fair use defense, and permits disproportionate statutory damages. While § 1202 includes an intent requirement, that requirement is so minimal that it fails to meaningfully limit the statute's reach—especially in the LLM context, where the act of violating § 1202 may itself satisfy the mental requirement.

In response to these concerns, this Note proposes that courts adopt an identicality requirement for § 1202 claims against LLMs. This requirement would cabin liability to outputs that exactly match training data—cases where the removal or alteration of CMI is both clear and technically avoidable. The approach mirrors Congress's existing accommodation of technological limitations for broadcasters in cases of technical infeasibility or financial hardship. A similar understanding for LLMs would preserve § 1202's core purpose, resolve a growing district court split, and ensure the most consequential copyright question raised by generative artificial intelligence is answered on its merits—not sidestepped by a statute never designed to decide it.

[†] [Acknowledgment Text]

Table of Contents

There are several differences between the standards for a § 1202(b) violation and copyright infringement. Although § 1202 was intended to adapt copyright enforcement to the challenges of the digital age, it has evolved into a powerful standalone doctrine that threatens to expand liability far beyond traditional infringement. Direct infringement is a strict liability offense[69] that explicitly requires copyrightable subject matter,[70] follows a two-part test for proving liability,[71] permits a fair use defense,[72] and limits damages on a per-work basis.[73] By contrast, § 1202 features a loosely applied scienter requirement,[74] no express copyrightability threshold,[75] no recognized fair use defense,[76] and damages that accrue on a per-violation basis.[77] In addition, whereas copyright registration is required to bring a copyright infringement claim, there is no such requirement for a § 1202 claim.[78] The result is that § 1202 has the potential to impose massive liability for removing CMI from minimally original works, for uses that would otherwise qualify as fair use, and for producing works that only loosely resemble the alleged source material.[79] This section examines the mechanisms of that expansion by comparing § 1202's scienter, copyrightability, liability threshold, defenses, and damages elements to those of traditional infringement doctrine.

A. Weakened Standard for Proving Copying

The standard for determining whether copying rises to liability in a copyright infringement claim consists of two parts: copying in fact and copying in law.

---

denying defendant's motion for summary judgment on plaintiff's DMCA claims). Note that Doe v. GitHub is an example of a case in which the plaintiffs allege violation of the DMCA without alleging copyright infringement. Doe v. GitHub, Second Amended Complaint at 49.

[69] *See infra* Part II.D.

[70] 17 U.S.C. § 102; *see infra* Part II.A.

[71] *See infra* Part II.A.

[72] *See infra* Part II.B.

[73] *See infra* Part II.C.

[74] *See infra* Part II.D.

[75] *See infra* Part II.A.

[76] *See infra* Part II.B.

[77] *See infra* Part II.C.

[78] *SellPoolSuppliesOnline.com LLC v. Ugly Pools Arizona, Inc.*, 344 F. Supp. 3d 1075, 1081 (D. Ariz. 2018) ("The text of the DMCA does not limit the protection of CMI to registered works."); *Shihab v. Complex Media, Inc.*, No. 21-CV-6425 (PKC), 2022 WL 3544149, at *8 (S.D.N.Y. Aug. 17, 2022) (concluding that 17 U.S.C. § 412's registration requirement does not apply to DMCA § 1202); Crusey, *supra* note 21 at 515 ("Unlike a copyright infringement claim under Section 501, a Section 1202 claim requires no prerequisite copyright registration ").

[79] Crusey, *supra* note 21 at 483 ("…[D]amages have the potential to accumulate to massive amounts in cases involving removal of CMI en masse, which results in many consecutive individual violations of Section 1202(b). In many of the recent AI copyright lawsuits, CMI is alleged to have been removed from millions of works, leading to potential Section 1202 damage awards in the billions of dollars.").

Copying in fact establishes that copying actually happened, and copying in law determines whether that copying was legally significant.[80] Unlike infringement, the § 1202 standard for determining whether a legally significant copy has been made depends on the court. Some courts implement an identicality standard, under which the allegedly violating work must be identical to the allegedly violated, original work.[81] Other courts implement a similarity standard instead, under which similarity, rather than identicality, is enough to determine that a "copy" was made.[82]

Infringement. To show copying in fact, a plaintiff bringing an infringement claim must first show that the defendant actually copied the allegedly infringed work.[83] This inquiry exists to ensure plaintiffs and defendants did not independently create their respective works, a finding that would invalidate the infringement claim.[84]

Because direct evidence of copying is rare,[85] to show copying in fact, plaintiffs often use circumstantial evidence, including some combination of proof of the defendant's access to the plaintiff's work and the similarity between the works.[86]

Striking similarity alone may also be enough to show copying.[87] This is a strict standard; the plaintiff must show identicality or similarity "such that, considering all the circumstances, there is no reasonable explanation . . . other than copying."[88] Whether a work is strikingly similar overlaps somewhat with the "thinness" (or here, "thickness") of the work's copyright:[89] "the more

[80] *See infra* notes 83-98 and accompanying text.

[81] *See infra* note 111.

[82] *See infra* note 131.

[83] Restatement of the Law, Copyright § 7.03 TD No 4 (2023) ("Proving copying in fact requires the plaintiff to prove some material was copied from the plaintiff's copyrighted work…").

[84] Restatement of the Law, *supra* note § 7.03 ("If the defendant proves that the allegedly infringing material was independently created rather than copied, directly or indirectly, from the plaintiff's copyrighted work, there was no copying in fact from the plaintiff's work and therefore no liability for copyright infringement.").

[85] Jonathan Trinidad Lira, *The End of the Inverse Ratio Rule in the Ninth Circuit and Its Effect on the Copyright System*, Seton Hall Student Works 1, 5 ("For the factual question of copying, there is rarely direct evidence of defendant actually copying the original work; thus, the plaintiff often proves actual copying by circumstantial evidence . . . .").

[86] Lydia Pallas Loren & Anthony Reese, *Proving Infringement: Burdens of Proof in Copyright Infringement Litigation*, 23 Lewis & Clark L. Rev. 621, 657 (2019); Restatement of the Law, *supra* note § 7.03.

[87] *Id.*

[88] Restatement of the Law, *supra* note § 7.03.

[89] *See infra* notes 96-98 and accompanying text.

inventive or expressive the duplicated content, the more likely the duplication resulted from copying."[90]

In addition to copying in fact, infringement claims must show "copying in law."[91] The role of this prong is to decide whether an established act of copying legally rises to the level of infringement.[92]

In particular, copying in law ensures that only copyrightable material can be found to be infringed[93] and that the copying of this material is "substantial."[94] To that end, in the Ninth Circuit, for example, the substantial similarity test involves filtering for protectable components in the plaintiff's work before conducting an objective comparison of the two works, and then subjectively determining whether the "total concept and feel" of the two works is similar.[95] While the normal standard to find infringement requires substantial similarity, "when a work is entitled to 'thin' protection, courts typically require 'virtual identity'" instead.[96] This means the validity or strength of a work's copyright protection moderates the level of similarity required for a finding of substantial similarity. That is, where copyright is at its thinnest, as in a factual compilation,[97] "liability hinges on entire duplication."[98]

Various doctrines govern whether material is copyrightable. The idea-expression dichotomy, for example, permits protection only for a work's

[90] Robert F. Helfing, Substantial Similarity and Junk Science: Reconstructing the Test of Copyright Infringement, 30 Fordham Intell. Prop. Media & Ent. L.J. 735, 763-64 (2020).

[91] Mark Edward Blankenship Jr., *Dead Frogs, Dissected Jokes, & Thin Copyright: Analyzing Copyrightable Elements and Legal Protection of Stand-Up Comedy*, 2 Florida Entertainment & Sports L. Rev. 117, 128 (2022).

[92] Restatement of the Law, *supra* note § 7.04 ("Proving improper appropriation requires the plaintiff to prove that sufficient protected expression was copied from the copyrighted work such that there is a substantial similarity between the alleged infringing use and the copyrighted work.").

[93] *See* Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991) ("As a constitutional matter, copyright protects only those constituent elements of a work that possess more than a *de minimis* quantum of creativity."); *see* Blankenship, *supra* note at 127 ("[Copying in law] considers whether the defendant copied a copyrightable expression in such a way that liability should follow."); *See* Nimmer on Copyright, *supra* note 25 § 13D.06 ("[T]he substantial similarity step of the infringement analysis focuses only on the protected elements of plaintiff's work that defendant copied.").

[94] Alan Latman, *"Probative Similarity" as Proof of Copying: Toward Dispelling Some Myths in Copyright Infringement*, 90 Columbia L. Rev. 1187, 1189 (1990) ("[N]ot only must [defendant] copy protected material, but also such protected material must be 'substantial.'").

[95] *Rentmeester v. Nike, Inc.*, 883 F.3d 1111, 1118 (9th Cir. 2018) (overruled on other grounds).

[96] *Advanta-STAR Auto. Rsch. Corp. of Am. v. Search Optics, LLC*, 672 F. Supp. 3d 1035, 1049 (S.D. Cal. 2023).

[97] Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991) ("This inevitably means that the copyright in a factual compilation is thin.").

[98] Nimmer On Copyright, *supra* note 25 § 13D.32 (adding that multiple circuits agree with this approach, and that "virtual identity is not the same as absolute identity.").

expressive elements—not its underlying ideas.[99] For instance, a book describing a "particular bookkeeping system" is itself copyrightable, but the method of bookkeeping it illustrates is not.[100] Similarly, facts are not protected,[101] though the "selection and arrangement" of facts in a compilation may be.[102] The merger doctrine holds that if there are few ways of expressing an idea, that expression "merges" with the idea and is unprotectable.[103] For example, a jewel-encrusted pin shaped like a bee is unprotected because there are limited ways to depict that idea.[104]

is Finally, works in the public domain are not protected by copyright.[105]

§ 1202. Section 1202's framework does not include an analysis of copying in fact or copying in law. The provision does, however, require that a violating work is a "copy" of the violated work.[106] This requirement comes from § 1202(c)'s definition of "copyright management information" as various types of "information conveyed in connection with copies … of a work."[107] In other words, a work creating § 1202 liability cannot be a *new* work, unrelated to the allegedly violated work. To ensure this standard is met, courts have developed two different standards that serve as proxies for whether the work in question is a copy of the original, allegedly violated work. [108]

---

[99] *Id.* § 2.03.

[100] Edward Samuels, *The Idea-Expression Dichotomy in Copyright Law*, 56 Tenn. L. Rev. 321, 326-27 (1989).

[101] Eldred v. Ashcroft, 537 U.S. 186 (2003) ("[C]opyright gives the holder no monopoly on any knowledge. A reader of an author's writing may make full use of any fact or idea she acquires from her reading.").

[102] Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991) ("As applied to a factual compilation, assuming the absence of original written expression, only the compiler's selection and arrangement may be protected; the raw facts may be copied at will.").

[103] William E. Greenspan, *Applications of the Idea/Expression Merger Doctrine in Copyright Infringement Cases*, 26 Bus. L. Rev. 1, 2 (1993).

[104] Herbert Rosenthal Jewelry Corp. v. Kalpakian, 446 F.2d 738, 742 (9th Cir. 1971). Instead of finding the expression entirely unprotectable, some courts instead subject it to "thin" copyright. *See* David G. Luettgen, *Functional Usefulness vs. Communicative Usefulness: Thin Copyright Protection for the Nonliteral Elements of Computer Programs*, 4 Tex. Intell. Prop. L.J. 233, 255 n. 118 (1996) ("*Feist* and *Bonito Boats* suggest that copyright law should never afford borderline factual and functional works more than 'thin' protection against wholesale appropriation of surface expression.").

[105] Faith O. Majekolagbe, *Public Domain and Access to Knowledge*, 31 J. Intell. Prop. L. 1, 2 (2024).

[106] 26 U.S.C. § 1202(c) (defining "copyright management information" as "information conveyed in connection with copies…"). Note that the provision again refers to copies in § 1202(b)(3), which specifies "copies of works" in its prohibition of distribution, import for distribution, and public performance of works.

[107] *Id.*.

[108] One understanding is that courts derive their identicality and similarity standards from the requirement that a work be a "copy." *See* Nimmer On Copyright, *supra* note 25 § 12A.10(C)(1) ("[§ 1202(b)(3)'s mention of 'copies of works'] therefore serves as the statutory hook to prevent nonsensical claims that plaintiff has a right to complain about