

The Enforceability of License Restrictions on Ingestion for Artificial Intelligence Training Purposes

Jonathan Band and Krista Cox
April 3, 2024

Many forms of artificial intelligence require the ingestion of vast quantities of online works subject to copyright protection. There are powerful arguments that this ingestion does not infringe copyright because it is permitted under the fair use right.¹ However, providers of content online typically make their material available under terms of service (“TOS”) or other agreements that prohibit the ingestion of the content. The question is whether such prohibitions are enforceable in the United States against libraries and academic researchers; in other words, whether contract terms can override fair use rights. The answer, unfortunately, is not straightforward. In many situations, it is unclear whether a contract has been formed in the first place because the licensee may not have sufficiently manifested assent to the terms of the contract. And even if the licensee manifested assent to the license terms, the terms may be preempted because they conflict with federal copyright law. In contrast to the unsettled state of the law in the United States, the European Union prohibits the enforcement of license restrictions on ingestion of content by libraries and research organizations for artificial intelligence purposes.

¹ See Library Copyright Alliance, Comments of the Library Copyright Alliance on the Inquiry Concerning Artificial Intelligence and Copyright (Oct. 2023), <https://www.librarycopyrightalliance.org/wp-content/uploads/2023/10/CO-AI-NOI-Final.pdf>.

A. Manifestation of Assent

A contract typically is enforceable against a party if the party manifested its assent to the contract's terms. The traditional way of manifesting assent is signing a contract. In the digital age, new forms contracts with new forms of manifesting assent have emerged. The software industry invented "shrink-wrap" licenses, which included statements on the plastic wrapping of boxes containing discs that provided that by removing the wrapping, users agreed to software license terms. While some manufacturers printed the full text of the license on the packaging, others included the full terms only inside the box. Later, "click-wrap" licenses were developed, where the user had to click on an "I agree" icon while installing software into a computer. By clicking on the icon, the user manifested her assent to the TOS. Similar "click-wrap" licenses are used on some websites; without clicking on the icon, the user cannot access the site.

Alternatively, some websites use employ a "browse-wrap" license, where the TOS declare that by using the website, the user agrees to the website's TOS. With a browse-wrap license, the user can access the website without expressly agreeing to the TOS; indeed, the user might not even know of the existence of the TOS. The website operator often places a link to the TOS at the bottom of the website's home page, and the user learns of the existence of the TOS only by affirmatively seeking them out.

There have been many cases examining the enforceability of shrink-wrap, click-wrap and browse-wrap licenses, and the results are far from uniform.² A central issue in

² Compare *Novell, Inc. v. Network Trade Ctr., Inc.*, 25 F. Supp. 2d 1218, 1230 (D. Utah 1997); *Morgan Labs., Inc. v. Micro Data Base Sys., Inc.* 41 U.S.P.Q.2d 1850 (N.D. Cal.

these cases is whether the user sufficiently manifested her assent to the license in order to be bound to its terms. As a general matter, courts tend to enforce click-wrap licenses, particularly against commercial users, because it is hard for the user to argue that it did not assent to the TOS—after all, it clicked on the “I agree” icon. Moreover, courts have little sympathy for the argument that a commercial user did not understand the provisions of the TOS—an argument that works better when the user is an individual consumer.

Courts have been less likely to enforce browse-wrap licenses because the user can more plausibly argue that it did not manifest assent to the TOS. Thus, in *Nguyen v. Barnes & Noble*, 763 F. 3d 1171 (9th Cir. 2014), the Ninth Circuit did not enforce an arbitration clause in a browse-wrap agreement where there was no notice to users or prompts to demonstrate assent, there was no evidence that the plaintiff was aware of the browse-wrap agreements terms. However, the Ninth Circuit noted, “courts have consistently enforced browsewrap agreements where the user had actual notice of the agreement.” *Id.* at 1176. Courts have also noted that constructive knowledge of the TOS in browse-wrap agreements may suffice. *Meyer v. Uber Technologies, Inc.*, 868 F.3d 66 (2d Cir. 2017) (“[T]he determination of the validity of the browsewrap contract depends on whether the user has actual or constructive knowledge of a website’s terms and

1997); *Arizona Retail Sys., Inc. v. The Software Link, Inc.*, 831 F. Supp. 759, 764-66 (D. Ariz. 1993); *Step-Saver Data Sys. v. Wyse Tech.*, 939 F.2d 91, 98-100 (3d Cir. 1991); *Foresight Resources Corp. v. Pfortmiller*, 719 F. Supp. 1006, 1010 (D. Kan. 1989); *Ticketmaster Corp. v. Tickets.com, Inc.*, 2000 U.S. Dist. LEXIS 12987 (C.D. Cal. August 10, 2000), *aff’d*, 248 F. 3d 1173 (9th Cir. 2001); *Specht v. Netscape Communications Corp.*, 306 F.3d 17 (2d Cir. 2002); *Softman Prods. Co. v. Adobe Systems, Inc.*, 171 F. Supp. 2d 1075 (C.D. Cal. 2001); and *Klocek v. Gateway, Inc.*, 104 F. Supp. 2d 1332, 1338-39 (D. Kan. 2000); *with ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1449 (7th Cir. 1996); *Register.com, Inc., v. Verio, Inc.*, 356 F.3d 393 (2d Cir. 2004); *cf. Hill v. Gateway 2000, Inc.*, 105 F.3d 1147, 1150 (7th Cir. 1997), *cert. denied*, 522 U.S. 808 (1997).

conditions . . . While it may be the case that many users will not bother reading the additional terms, that is the choice the user makes; the user is still on inquiry notice.”)

Given the absence of a clear manifestation of assent with browse-wrap licenses, website operators know that if they want to ensure that courts will enforce their TOS, they should use a click-wrap license. However, website operators often make the business decision not to employ a click-wrap license. They know that users may be unwilling to click on them, and instead might proceed to a competing website that does not require clicking on an “I agree” icon.

Notwithstanding judicial reticence to enforce browse-wrap licenses, library personnel may have actual knowledge of browse-wrap prohibitions on the downloading of content for AI ingestion purposes, or a court might find that library personnel had constructive knowledge of such terms. A research library certainly would be found to have knowledge of the terms in an e-book or e-journal license agreement with a publisher. Under such circumstances, federal preemption of enforcement of contract terms, discussed in the next section, becomes relevant.

B. Preemption

Courts might refuse to enforce contractual restrictions on downloading content from a website because such restrictions conflict with the U.S. Copyright Act’s fair use right codified at 17 U.S.C. § 107. Under certain circumstances, federal copyright law may preempt state contract law. The law in this area, however, is far from certain.

Section 301(a) of the Copyright Act, 17 U.S.C. § 301(a), provides that the Copyright Act preempts state rights “that are equivalent to any of the exclusive rights within the general scope of copyright....” Courts have interpreted Section 301(a) as not

preempting a state cause of action that requires proof of “extra elements” not present in a copyright claim. The Seventh Circuit in *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447 (7th Cir. 1996), ruled that Section 301(a) did not preempt enforcement of a contract prohibiting the copying of telephone listings because the contract claim required proof of an extra element—the existence of an enforceable contract.

However, in *Data General Corp. v. Grumman Systems Support Corp.*, 36 F.3d 1147 (1st Cir. 1994), the First Circuit noted that not every extra element will establish a qualitative variance between rights under copyright and those protected by state law. Thus, if the extra elements are “illusory ... mere labels attached to the same odious business conduct,” then preemption will occur. Likewise, a leading treatise on copyright law endorses the position that preemption should be found to the extent that a breach of contract cause of action alleges nothing more than simply reproduction (or adaptation, distribution, etc.) of a copyrighted work. 1 Nimmer on Copyright § 1.01[B][1][a][i]. Nimmer argues that causes of action denominated as contract claims “at times may essentially allege nothing other than derogation of rights under copyright.” *Id.* In those cases, “a breach of contract cause of action can serve as a subterfuge to control nothing other than the reproduction, adaptation, public distribution, etc. of works within the subject matter of copyright. Those instances are to be deemed pre-empted.” *Id.* § 1.01[B][1][a][iii]. Numerous courts have relied upon Nimmer’s reasoning.

The district court in *Canal+ Image UK Ltd. v. Lutvak*, 773 F.Supp.2d 419 (S.D.N.Y. 2011), for example, stated that Nimmer’s position made sense “because a rule that ‘the extra element that saves a contract claim from preemption is the promise itself,’ ... seems hard to square with the Second Circuit’s instruction to ‘take a restrictive view

of what extra elements transform an otherwise equivalent claim into one that is qualitatively different from a copyright infringement claim.”³

Courts have also preempted contractual terms on Constitutional grounds. In 1988, the U.S. Court of Appeals for the Fifth Circuit set aside a contractual restriction on reverse engineering in *Vault Corp. v. Quaid Software Ltd*, 847 F.2d 255 (5th Cir. 1988). The *Vault* court cited *Sears, Roebuck & Co. v. Stiffel Co.*, 376 U.S. 225, 229 (1964), where the Supreme Court relied on the U.S. Constitution’s Supremacy Clause to conclude that “[w]hen state law touches upon an area of [the copyright statutes], it is ‘familiar doctrine’ that the federal policy ‘may not be set at naught, or its benefits denied’ by state law.” The *Vault* court held that a reverse engineering prohibition in a shrink-wrap license “conflicts with the rights of computer program owners under Section 117 and clearly ‘touches upon an area’ of federal copyright law.” *Vault*, 847 F.2d at 270. Likewise, the Supreme Court relied on the Supremacy Clause to preempt a Florida plug mold statute it found inconsistent with the federal intellectual property system. *Bonito Boats Inc., v. Thunder Craft Boats Inc.*, 489 U.S. 141 (1989).

On the other hand, the courts in *Bowers v. Baystate Techs., Inc.*, 320 F.3d 1317 (Fed. Cir. 2003), *cert. denied*, 539 U.S. 928 (2003), and *Davidson & Assoc. v. Jung*, 422 F.3d 630 (8th Cir. 2005), rejected constitutional preemption arguments with respect to contractual restrictions on copyright exceptions. Judge Dyk, however, wrote a powerful

³ See also *ML Genius Holdings LLC v. Google LLC*, No. 20-3113 (2d. Cir. judgment entered March 31, 2022); *Selby v. New Line Cinema*, 96 F. Supp. 2d 1053 (C.D. Cal. 2000); *Symantec Corp. v. McAfee Associates*, 1998 WL 740798 (N.D. Cal. June 9, 1998); *Ass’n for Info. Media and Equip. v. Regents of the Univ. of California*, No. 2:10-CV-09378-CBM, 2012 WL 7683452 (C.D. Cal. Nov. 20, 2012); *Kabehie et al., v. Zoland, et al.*, 125 Cal. Rptr. 721 (Cal.App.2nd Dist. 2002).

dissent in *Bowers*, stating that a software firm could not eliminate a user’s privileges under the Copyright Act simply “by printing a few words on the outside of its product....” *Bowers*, 320 F.3d at 1337 (Dyk, J., dissenting). Such an approach “permits state law to eviscerate an important federal copyright policy reflected in the fair use defense....” *Id.* at 1335.

C. Conclusion

A library might be able to make a credible argument that it did not manifest assent to a browse-wrap prohibition on downloading content from a website for AI ingestion purposes. It would have greater difficulty making this argument in the context of an e-book or e-journal license agreement. Nonetheless, there is case law supporting a library’s argument that enforcement of such a contractual restriction is preempted by federal law. Unfortunately, the law in this area is far from clear and requires greater development.

It should be noted that the European Union has resolved this question in favor of libraries. Article 3 of the Copyright in the Digital Single Market Directive requires Member States to enact copyright exceptions permitting “reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining⁴ of works or other subject matter to which they have lawful access.” Article 7(1) of the Directive stipulates that “[a]ny contractual provision contrary to the exception[] provided for in Article[] 3... shall be unenforceable.” Thus, in the EU, a publisher could not enforce a license term prohibiting

⁴ Article 2(2) defines “text and data mining” as “any automated technique aimed at analyzing text and data in digital form in order to generate information which includes but is not limited to patterns, trends, and correlations.” This definition is broad enough to include ingestion for AI training purposes.

a research organization or cultural heritage institution from ingesting the licensed material for AI ingestion purposes.⁵

Because Congress is unlikely to enact this sort of legislative relief, U.S. libraries must attempt to include fair use savings clauses in their licenses with publishers to ensure that the licenses do not curtail their fair use rights with respect to ingestion for AI training purposes.

⁵ Significantly, the Article 7(1) prohibition on contracts overriding the Article 3 text and data mining exception does not apply to the Article 4 text and data mining exception for all entities *other* than research organizations and cultural heritage institutions. So publishers can employ license terms to restrict the effectiveness of the Article 4 TDM exception. For a more detailed discussion of the AI-related exceptions in foreign jurisdictions, see Jonathan Band, *Copyright Implications of the Relationship between Generative Artificial Intelligence and Text and Data Mining*, infojustice (Oct. 2023), <https://infojustice.org/archives/45509>.