



COMMENTS OF THE LIBRARY COPYRIGHT ALLIANCE ON REMOVING BARRIERS TO AMERICAN LEADERSHIP IN ARTIFICIAL INTELLIGENCE

This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in developing the AI Action Plan and associated documents without attribution.

The Library Copyright Alliance (“LCA”) consists of the American Library Association and the Association of Research Libraries. These two library associations represent over 100,000 libraries in the United States employing over 300,000 librarians and other personnel.

LCA appreciates the opportunity to respond to the Networking and Information Technology Research and Development National Coordination Office, National Science Foundation, request for input on the development of an Artificial Intelligence (“AI”) Action Plan directed towards sustaining and enhancing America’s AI dominance. We believe the highest priority action items relate to two dimensions of ensuring that copyright law does not become a barrier to the development of AI:

- First, the courts should treat the ingestion of content necessary to power many AI activities as a fair use under 17 U.S.C. § 107. To that end, the Action Plan should call for the Administration to file *amicus* briefs in support of fair use in appropriate cases.
- Second, license terms purporting to abridge the fair use rights of users should not be enforced. This can be achieved under theories such as conflict preemption advocated by the Administration in *amicus* briefs in litigation.

These issues are discussed below in greater detail.

I. Preserving Fair Use

American leadership in AI is attributable in large measure to the fair use right, codified at 17 U.S.C. § 107. For more than two decades, both academic and commercial AI researchers have swept up enormous amounts of content from disparate sources, including the Internet, in the course of developing AI processes. Most of this content is protected by copyright, and the amount of content needed for AI training is so vast that identifying all the rightsholders and receiving their permission is a practical impossibility. Fortunately for U.S. AI researchers, a

series of appellate fair use decisions starting in the early 1990s provided a legal basis for this aggregation without the rightsholders' authorization.¹

In contrast, the UK Copyright, Design, and Patent Act was amended only in 2014 to permit the copying incidental to computational analysis, and then just for non-commercial purposes. Copying for computational analysis was allowed in the EU and Japan in 2019. This meant that the legal environment in the United States provided U.S. AI researchers as much as a two-decade lead time advantage over competitors in Europe and Japan.

To be sure, the U.S. appellate decisions referenced above did not concern *generative* AI; but the reasoning of the courts in these cases weigh heavily in favor of a fair use finding in the pending cases concerning generative AI. Nonetheless, some of the preliminary district court decisions in the recent surge of AI cases² suggest that the courts could benefit from the Administration's views on the fair use nature of the copying incidental to the training of AI models.³ In appropriate cases, the Administration should file *amicus* briefs stressing the enormous public benefits of AI to the United States, and how these benefits greatly outweigh the speculative harm caused to copyright owners by the act of ingestion. The Supreme Court in *Google v. Oracle*, 141 S. Ct. 1183, 1206 (2021), underscored that the fair use calculus often "must take into account the public benefits the copying" under consideration is likely to produce. The Administration must make clear to the courts the public benefit of continued American leadership in AI.⁴

¹ *Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992) and *Sony Computer Entertainment, Inc. v. Connectix Corp.*, 203 F.3d 596 (9th Cir. 2000) permitted the copying of entire computer programs in the course of reverse engineering for the purpose of extracting the information necessary to achieve interoperability. Following these precedents, *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003), *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007), *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630 (4th Cir. 2009), *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014), and *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015), all approved reproductions during the course of various artificial intelligence processes.

² Ali Sternburg, *Scholars Agree Opinion in Thomson-Reuters v. Ross Should be Disregarded*, Disruptive Competition (Feb. 28, 2025), <https://project-disco.org/intellectual-property/scholars-agree-opinion-in-thomson-reuters-v-ross-should-be-disregarded/>.

³ The Administration should, in appropriate cases, also file *amicus* briefs arguing that the removal of copyright management information during the AI training process does not violate 17 U.S.C. § 1202.

⁴ In general, the Administration should take a more favorable view toward fair use than previous Administrations. In 2014, the revenue from industries reliant on fair use was \$5.6 trillion. The value added by these industries to the U.S. economy was over \$2.8 trillion. Computer & Communications Industry Association, *Fair Use in the U.S. Economy* (2017), <https://ccianet.org/wp-content/uploads/2017/06/Fair-Use-in-the-U.S.-Economy-2017.pdf>. These numbers are significantly larger now.

In the unlikely event that the courts apply fair use so narrowly that the development of generative AI is constrained, the Administration should urge Congress to enact legislation to enable the advancement of this technology. As noted, other countries with sophisticated technology sectors have enacted provisions in their copyright laws that facilitate AI. The United States must not place itself at a competitive disadvantage relative to other countries by hindering the evolution of this important tool. But at this point, we should let the courts—with the Administration’s guidance—apply fair use principles to the many different fact patterns presented in the cases centering on ingestion.

Although the ingestion of works for AI training purposes generally is a fair use, the copyright status of the AI’s outputs must be considered separately. If an AI produces a work that is substantially similar in protected expression to a work that was ingested by the AI, that new work likely infringes the copyright in the original work.⁵

II. Limiting the Adverse Impact of License Terms

While the appellate court decisions demonstrate that the ingestion of copyrighted materials for the purpose of an AI process is a fair use, the terms of click-wrap⁶ and browse-wrap⁷ license terms imposed by database providers and website operators could prohibit the ingestion of material. With browsewrap licenses, there are serious questions whether there was sufficient manifestation of assent to form a binding contract.⁸ Moreover, strong arguments can be made that browsewrap and clickwrap licenses inconsistent with the fair use right are preempted either by the Constitution⁹ or 17 U.S.C. § 301(a).¹⁰

Notwithstanding these arguments, there is contrary authority.¹¹ When this issue arises again in the courts, the Administration should file *amicus* briefs arguing that contractual

⁵ www.librarycopyrightalliance.org/wp-content/uploads/2023/06/AI-principles.pdf.

⁶ In a click-wrap license, the user must click on an “I agree” icon in order to access digital material such as a database or a website.

⁷ In a browse-wrap license, a website’s terms of service declare that by using the website, the user agrees to the website’s terms of service.

⁸ See *Nguyen v. Barnes & Noble*, 763 F.3d 1171 (9th Cir. 2014). See generally Jonathan Band and Krista Cox, *The Enforceability of License Restrictions on Ingestion for Artificial Intelligence Training Purposes*, available at arll.org (April 2024).

⁹ See *Vault Corp. v. Quaid Software Ltd*, 847 F.2d 255 (5th Cir. 1988); *X Corp. v. Bright Data Ltd.*, (N.D. Cal. May 9, 2024).

¹⁰ See *Data General Corp. v. Grumman Systems Support Corp.*, 36 F.3d 1147 (1st Cir. 1994); *ML Genius Holdings LLC v. Google LLC*, No. 20-3113 (2d Cir. Mar. 10, 2022); 1 Nimmer on Copyright § 1.01[B][1][a][i].

¹¹ See *Meyer v. Uber Technologies, Inc.*, 868 F.3d 66 (2d Cir. 2017); *Bowers v. Baystate Techs., Inc.*, 320 F.3d 1317 (Fed. Cir. 2003), *cert. denied*, 539 U.S. 928 (2003); *Davidson & Assoc. v. Jung*, 422 F.3d 630 (8th Cir. 2005); *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447 (7th Cir. 1996).

prohibitions on copying otherwise permitted by fair use are not enforceable.¹² If the courts ultimately get this issue wrong, and rule in favor of rightsholders unilaterally eviscerating fair use via contract, the Administration should urge Congress to enact legislation making clear that exceptions and limitations provided in the Copyright Act prevail over license terms inconsistent with those exceptions. It should be noted that for over thirty years, the European Union has included contract preemption clauses in its directives.¹³ It has recognized that it would be pointless to require Member States to adopt exceptions if private parties could simply override them by contract. Significantly, Article 7(1) of the Digital Single Market (“DSM”) Directive voids any contractual provision contrary to Article 3, which permits text and data mining by research organizations and cultural heritage institutions for scientific research.

III. Conclusion

To sustain and enhance America’s current AI dominance, U.S. courts must continue to find the ingestion of content needed for AI processes to be a fair use. To assist the courts achieving this result, the AI Action Plan should state that the Administration will file *amicus* briefs in appropriate cases explaining the public benefit of AI. Likewise, the AI Action Plan should aver that the Administration will file *amicus* briefs articulating the position that licenses prohibiting ingestion for AI purposes should not be permitted to trump the fair use right.

Respectfully submitted,

Jonathan Band
LCA Counsel
jband@policybandwidth.com

March 14, 2025

¹² This issue is different from bot exclusion headers, such as robot.txt, which website operators can employ to prevent their sites from being crawled for AI training purposes. Most AI firms respect these bot exclusion headers voluntarily.

¹³ See Jonathan Band, *Protecting User Rights Against Contract Override*, PIJIP/TLS Research Paper Series (May 2023) <https://digitalcommons.wcl.american.edu/research/97>.