COMMENTS OF THE LIBRARY COPYRIGHT ALLIANCE ON INTELLECTUAL PROPERTY PROTECTION FOR ARTIFICIAL INTELLEGENCE INNOVATION

The Library Copyright Alliance ("LCA") consists of the American Library Association, the Association of College and Research Libraries, and the Association of Research Libraries. Collectively, these three library associations represent over 100,000 libraries in the United States employing over 350,000 librarians and other personnel.

LCA appreciates the opportunity to respond to the USTPTO's request for comments regarding intellectual property protection for artificial intelligence ("AI") innovation published in the Federal Register at 84 Fed. Reg. 58141 on October 30, 2019. LCA's response will focus on question 3: whether the statutory language of the fair use doctrine and related case law adequately address the lawfulness of the ingesting of large volumes of copyrighted material necessary for an AI algorithm or function. LCA believes that the fair use right is adequate to this task. However, license terms employed by website operators and database providers could interfere with the ingestion of materials. A statutory "contract override" provision similar to that found in various EU directives may be necessary to resolve this problem.

I. Fair Use

AI is opening new fields of scholarly research such as the digital humanities. *See* Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. Copyr. Soc. USA __ (forthcoming 2019); Michael Carroll, *Copyright and the Progress of Science: Why Text and Datamining is Lawful*, 53 U.C. Davis L. Rev. __ (forthcoming 2019). This research requires the creation of searchable databases where AI is employed to enable users to detect patterns across a large number of works.

Recent fair use jurisprudence makes clear that the copying necessary to performed an AI process is a fair use. One of the leading cases is *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014). The HathiTrust Digital Library ("HDL") contains electronic copies of more than ten million books in HathiTrust members' collections digitized by Google in the course of the Google Books Project.[1] The Authors Guild sued for declaratory judgment and injunctive relief. Both the district court and the Second Circuit found that the copies made by HathiTrust were permitted by the fair use right, 17 U.S.C. § 107.

[1] HDL created and maintained four copies of its entire database (one on the primary server at the University of Michigan, another at the mirror server at the Indiana University, and two encrypted back up tapes at two secure locations on the University of Michigan campus). 755 F.3d at 92. The copy of each work contains the full text of the work in machine readable format, as well as images of each page of the work as they appear in the print version. Thus, HDL holds eight permanent copies of each work. *Id*. In addition to preserving the books in the repository, HDL enables full-text search of the books and provides full text access to people with print disabilities.

In finding that HathiTrust's provision of full-text search functionality was a fair use, the court relied heavily on the Court's decision in *Campbell v. Acuff-Rose Music, Inc.,* 510 U.S. 569 (1994), particularly *Campbell'*s focus on the importance of transformative use under the first fair use factor, the purpose and character of the use. The *HathiTrust* court concluded that "creation of a full-text searchable database is a quintessentially transformative use," because it "does not 'supercede the objects or purposes of the original creation,'" 755 F.3d at 97 (quoting *Campbell*, 510 U.S. at 579). The *HathiTrust* court also relied on two Ninth Circuit decisions concerning Internet search engines—*Kelly v. Arriba Soft Corp*., 336 F.3d 811 (9th Cir. 2003) and *Perfect 10, Inc., v. Amazon.com, Inc*., 508 F.3d 1146 (9th Cir. 2007)—as well as a Fourth Circuit decision involving a plagiarism detection database—*A.V. ex rel. Vanderheye v. iParadigms, LLC,* 562 F.3d 630 (4th Cir. 2009). A year after *HathiTrust*, another Second Circuit panel reaffirmed *HathiTrust*'s holdings in *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015).

II. Contracts

While this cluster of circuit court decisions demonstrates that the ingestion of copyrighted materials for the purpose of an AI process is a fair use, the terms of click-wrap[2] and browse-wrap[3] 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.[4] Moreover, strong arguments can be made that browsewrap and clickwrap licenses inconsistent with the fair use right are preempted either by the Constitution[5] or 17 U.S.C. § 301(a).[6]

Notwithstanding these arguments, there is authority suggesting that contractual prohibitions on copying otherwise permitted by fair use may be enforceable.[7] Congressional intervention may be necessary to make clear that exceptions and limitations provided in the Copyright Act prevail over license terms inconsistent with those exceptions.

It should be noted that for nearly thirty years, the European Union has included contract preemption clauses in its directives. It has recognized that it would be pointless to require Member States to adopt exceptions if private parties could simply override them by contract.

- **Software Directive (1991).** Article 9(1) of the Software Directive provides that "[a]ny contractual provisions contrary to Article 6 or to the exceptions provided for in Article 5(2) and (3) shall be null and void." Article 5(2) permits back-up copying; Article 5(3)

[2] 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.
[3] 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.
[4] *See Nguyen v. Barnes & Noble*, 763 F. 3d 1171 (9th Cir. 2014).
[5] *See Vault Corp. v. Quaid Software Ltd*, 847 F.2d 255 (5th Cir. 1988).
[6] *See Data General Corp. v. Grumman Sytems Support Corp.*, 36 F.3d 1147 (1st Cir. 1994); 1 Nimmer on Copyright § 1.01[B][1][a][i].
[7] *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. Zeidenber*g, 86 F.3d 1447 (7th Cir. 1996).

permits "black box" reverse engineering; and Article 6 permits decompilation for the purpose of achieving interoperability.

- **Database Directive (1996).** Article 15 provides that "[a]ny contractual provision contrary to Articles 6(1) and 8 shall be null and void." Article 6(1) permits "acts necessary for the purpose of access to and normal uses of contents of database;" Article 8 permits a database user to extract and re-utilize insubstantial parts of a database.

- **Marrakesh Directive (2013).** Article 3 of **t**he Marrakesh Directive provides that "Member States shall ensure that the exception provided for in paragraph 1 cannot be overridden by contract." Paragraph 1 is the operative part of the Directive that permits authorized entities to make and distribute accessible format copies.

- **Digital Single Market Directive (2019).** Article 7(1) of the Digital Single Market ("DSM") Directive provides that "Any contractual provision contrary to the exceptions provided for in Articles 3, 5 and 6 shall be unenforceable." Article 3 permits text and data mining by research organizations and cultural heritage institutions for scientific research; Article 5 permits the use of works in digital and cross-border teaching activities; and Article 6 permits preservation by cultural heritage institutions.

All EU Member States must implement these contract preemption provisions in their law. Moreover, some Member States have adopted more extensive contact preemption provisions than those required by EU directives. For example, in addition to the contract preemption provisions required by the EU directives, the United Kingdom had declared unenforceable a term of a contract purporting to prevent the making of a copy which does not infringe copyright by virtue of exceptions in sections: 28B (personal use), 29 (research and private study), 30 (criticism, review, quotation, and news reporting), 31 (caricature and parody), 32 (illustration for education), 41(5) (supply of copies to other libraries), 42(7) (replacement copies), and 42A (single copy to user). The copyright laws of Germany, Ireland, Portugal, Montenegro, and Belgium likewise prevent the enforcement of contractual provisions restricting activities permitted by a wide range of exceptions.[8]

As noted above, the DSM Directive provides that any contractual restriction on the mandatory exception for text and data mining by research organizations be unenforceable. USPTO should consider proposing a similar provision that preempts any contractual restriction on the ingestion of material for AI processes.

Jonathan Band
LCA Counsel
jband@policybandwidth.com

January 10, 2020

[8] *See* International Federation of Library Associations and Institutions, *Protecting Exceptions Against Contractual Override*, https://www.ifla.org/files/assets/hq/topics/exceptions-limitations/documents/contract_override_article.pdf.