BEFORE THE HOUSE COMMITTEE ON THE JUDICIARY
SUBCOMMITTEE ON COURTS, INTELLECTUAL PROPERTY AND THE INTERNET

HEARING ON PRESERVATION AND REUSE OF COPYRIGHTED WORKS

STATEMENT OF JAMES G. NEAL
VICE PRESIDENT FOR INFORMATION SERVICES AND UNIVERSITY LIBRARIAN
COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Chairman Coble, Ranking Member Nadler, Members of the Subcommittee, my name is James G. Neal and I am the Vice President for Information Services and University Librarian at Columbia University in the City of New York. My testimony is endorsed by the Library Copyright Alliance (LCA).[1]

I appreciate the opportunity to testify today on the important issue of the preservation and reuse of copyrighted works.[2] In this statement, I will address four issues. First, I will describe the importance of library preservation with some examples of the preservation activities at Columbia. Second, I will explain how the library exceptions in Section 108 of the Copyright Act supplement, and do not supplant, the fair use right under Section 107, for important library activities such as preservation. Third, I will discuss changes in the legal landscape that diminish the need for legislation concerning orphan works. Fourth, I will provide the subcommittee my perspective on the *HathiTrust* case. My overarching point is that the existing statutory framework, which combines the

[1] LCA consists of three major library associations—the Association of Research Libraries, the American Library Association, and the Association of College and Research Libraries—that collectively represent over 100,000 libraries in the United States employing over 350,000 librarians and other personnel.

[2] This January, the Subcommittee held a hearing on the scope of fair use. In connection to that hearing, LCA submitted a statement for the hearing record on how fair use was integral to the ability of all types of libraries to achieve many facets of their mission, including preservation. The statement specifically discussed the importance of fair use for mass digitization, access to orphan works, and access to users with print disabilities. The statement explained how the library community had developed a code of best practices for the application of fair use by librarians. And the statement noted that recent judicial decisions, such as *Authors Guild v. HathiTrust*, had confirmed our understanding of how our practices were consistent with fair use.

specific library exceptions in Section 108 with the flexible fair use right, works well for libraries, and does not require amendment.

But before diving into copyright law, I would like to make clear to the subcommittee that libraries are not seeking a free ride. U.S. libraries spend over $4 billion a year acquiring books, films, sounds recordings, and other materials. Our objective is to maximize the benefit the American people receive from this enormous investment that ultimately they themselves make, by funding libraries through taxes or tuition, in order to purchase this material. We want to make sure that this material is accessible to the current generation of users, and that it is preserved so that it can be used by future generations. Libraries think in terms of centuries, not quarterly royalty reports. For almost four hundred years, libraries in America have promoted culture and democratic values, and we intend to continue doing so indefinitely.

I. THE IMPORTANCE OF THE LIBRARY PRESERVATION FUNCTION.

Libraries provide access to their collections of preserved materials as an essential component of their mission. Libraries engage in preservation activities to prevent the loss of vital cultural, historical and scholarly resources. A vast amount of material lacks commercial value or the publisher may not have the interest, financial incentive or technical expertise to engage in preservation activities. It is important to note that the amount of materials demanding preservation far exceeds the capacity of cultural memory organizations to fund, organize, and curate collections, forcing these organizations to make hard, technical decisions which materials to preserve.

The nature of library collections is changing and with change, come new challenges for preservation. Paper-based books and manuscripts have been the mainstay of scholarly communications and library collections for hundreds of years. But in less than two decades, digital information has become integral to research in all disciplines and to the public. Web documents, moving images, sound recordings, and data sets are an increasing part of everyday life and communication for much of the world. Rapidly these media are forming a substantial part of our cultural record some of which libraries are preserving locally or collaboratively through partnerships, consortia and new initiatives such as the HathiTrust and the Digital Preservation Network (DPN). One need only consider recent advances of digital technologies to understand that the preservation of

materials is necessary. Websites come and go, documents disappear from websites, hyperlinks get broken, files become corrupted and storage media become obsolete.

At Columbia, there are a significant number of collections that demand preservation, which may include shifting formats as some formats become obsolete. For example, the 9/11 Oral History Project focuses on the aftermath of the destruction of the World Trade Center. The Project amounts to over 900 recorded hours, including 23 hours on video with over 600 individuals—all recorded on digital media. The collection includes over 500 minidiscs, DAT tapes, and other media, recorded in 2002-10 and consisting of oral histories with people from a wide variety of ethnic and religious backgrounds involved with the 9/11 tragedy, including survivors, first responders, and people who lost friends and family members. Minidiscs were a short-lived medium that is now inaccessible due to the disappearance of the players. DAT tape deteriorates rapidly. More than half of this collection is already open and available to the public at Columbia, and the entire archive will, in due course, be available for study and research. This is only one of hundreds of such projects within the Columbia Center for Oral History, founded in 1948 and one of the largest oral history archives in the world.

Another example is the Language and Culture Archive of Ashkenazic Jewry, which includes over 5,700 hours of interviews mostly with surviving European Yiddish-speaking informants, collected between 1959 and 1972 in various countries on 2,552 reels of tape. While the purpose of the interviews was linguistic documentation, they include information about pre-World War II customs, culture, and experiences. Without the help of the National Endowment for the Humanities, New York State, and several private foundations who funded the preservation effort, the audiotapes would still be deteriorating and inaccessible.

Finally, the Human Right Archive, begun in 20008, is an innovative approach to documenting the state and progress of human rights around the world. Columbia is making complete copies, on a quarterly basis, of more than 600 websites from around the world, including sites covering human rights in Africa, Asia, the Middle East and South America. The archive now consists of more than 60 million pages, including many short-lived websites from countries in conflict or with repressive governments. This archive contains unique material that may in some cases be the only surviving records of regional

and citizen-based human rights organizations in countries like Uganda, Tibet, Ukraine and Venezuela. Columbia is creating a number of other targeted web archives, all bringing with them the need for long-term digital preservation.

In short, digital resources are not immortal. In fact, they are in formats that are more likely to cease to exist, and must be transferred to new digital formats repeatedly as technology evolves. They require extensive, highly specialized preservation and curation using constantly evolving methods and technologies. This means that the libraries charged with this work require robust applications of flexible exceptions such as fair use so that copyright technicalities do not interfere with their preservation mission.

II. THE RELATIONSHIP BETWEEN SECTION 108 AND FAIR USE.

A. Section 108 and the Privileged Status of Libraries in the Copyright Act.

Recognizing the importance of libraries to American democracy and culture, Congress has accorded them privileged status in Title 17. Section 109(b)(2) excludes libraries from the prohibition on software rental. Section 504(c)(2) shields libraries from statutory damages liability where they reasonably (but incorrectly) believed their actions constituted fair use. Section 602(a)(3)(C) provides organizations operated for scholarly, educational, or religious purposes with an exception to the importation right for "library lending or archival purposes." Section 1201(d) of the Digital Millennium Copyright Act ("DMCA") gives libraries the right to circumvent technological protection measures for purposes of determining whether to acquire a copy of the work. Section 1203(c)(5)(B) allows a court to remit statutory damages to libraries in cases of innocent violations of the DMCA. Section 1204(b) excludes libraries from criminal liability for DMCA violations.

More significantly, Congress enacted Section 108 in 1976 to provide libraries and archives with a set of clear exceptions with regard to the preservation of unpublished works; the reproduction of published works for the purpose of replacing a copy that was damaged, deteriorating, lost, or stolen; and the making of a copy that would become the property of a user.

Over the past 38 years, Section 108 has proven essential to the operation of libraries. It has guided two core library functions: preservation and inter-library loans. In the ongoing litigation between the Authors Guild and HathiTrust, however, the Authors Guild has attempted to convert a very helpful exception adopted to benefit

libraries into a limitation on the operation of libraries. This flawed interpretation of Section 108 harms libraries and departs from the plain language of the statute.

B. Section 108 Does Not Constrain Library Practices.

In the District Court, the Authors Guild argued that the Section 108 library exceptions represented the totality of the exceptions to the reproduction and distribution rights available to libraries. Under the Author Guilds' original position, libraries could not employ the first sale doctrine to circulate books, *see Kirtsaeng v. John Wiley & Sons, Inc.*, 133 S. Ct. 1351, 1364 (2013); nor Section 117(a) to copy software into their computers' memory; nor Section 109(c) to display book covers and posters in exhibitions; nor Section 110(1) to perform films in classrooms; nor Section 110(2) to perform and display works in distance education; nor Section 121 to make and distribute copies in accessible formats. Further, the Authors Guild argued that Section 108 precluded libraries from asserting the fair use right. Judge Baer correctly rejected these assertions.

In the Second Circuit, the Authors Guild more narrowly argued that HathiTrust "exceeded many of the express limitations of Section 108, and these violations should weigh heavily against a finding of fair use." As noted above, libraries rely on fair use to engage in a wide range of activities not covered by Section 108. If the Authors Guild's position were correct, libraries across the country would likely infringe copyright millions of times every day. For example, a major function of public libraries is providing free Internet access. Whenever a user views a website, the browser caches a copy of the website in the computer's memory. Courts have treated this cache copy as a fair use. *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1169-70 (9th Cir. 2007). Librarians and library users make hundreds of thousands, if not millions, of such copies every day. Because the cache copying by libraries in the course of Internet browsing wildly exceeds that authorized by Section 108, according to the Authors Guild, "these violations should weigh heavily against a finding of fair use."

Libraries regularly rely upon fair use to perform a wide range of other completely non-controversial practices. Libraries make preservation copies of musical, pictorial, graphic or sculptural works, and motion pictures—all categories of works not covered by Section 108. *See* 17 U.S.C. § 108(i); ARL, *Code of Best Practices in Fair Use for*

Academic and Research Libraries 17-18 (2012) ("*Code*"). Libraries archive websites of significant cultural or historical interest. *Code* at 26. They reproduce selections from collection materials to publicize their activities or to create physical and virtual exhibitions. *Id.* at 15. Academic libraries copy material into institutional digital repositories and make deposited works publicly available. *Id.* at 23. School libraries make multiple copies of appropriate portions of works for classroom use.

The Library of Congress, where the Copyright Office resides, relies heavily on fair use. For numerous collections, the Library of Congress states that it is providing online access to items "under an assertion of fair use" if "despite extensive research, the Library has been unable to identify" the rightsholder. *E.g.*, Library of Congress, *Copyright and Other Restrictions*, Prosperity and Thrift, http://memory.loc.gov/ammem/coolhtml/ccres.html. Similar language appears on the copyright pages of more than a dozen other collections. Under the Authors Guild' interpretation of Section 108, the Library of Congress is a serial copyright infringer.[3]

C. Section 108(f)(4) Unambiguously Provides that Section 108 Does Not Limit The Applicability of Fair Use to Libraries.

The plain language of Section 108, and its legislative history, underscore that Section 108 does not limit the availability of fair use to libraries. When the Senate Subcommittee on Patents, Trademarks, and Copyrights reported out the bill in December 1969 with the basic elements of what is currently Section 108, it included the language now in Section 108(f)(4). That clause provides that nothing in Section 108 "in any way affects the right of fair use as provided by section 17…." The Subcommittee report's discussion of Section 108 stated: "[t]he rights given to the libraries and archives by this provision of the bill are in addition to those granted under the fair-use doctrine." S. Rep. No. 91-1219, at 6 (1970). Section 108(f)(4) was added to address concerns some in the library community had raised about the potential impact of Section 108 on fair use.

The House Judiciary Committee Report on the 1976 Act quoted the language of Section 108(f)(4) and then explained that "[n]o provision of section 108 is intended to

[3] Other federal libraries also rely on fair use. *E.g.*, Smithsonian Institution Libraries, *Proceedings of the National Academy of Sciences of the United States*, Electronic Resources from the Smithsonian Libraries, http://www.sil.si.edu/eresources/silpurl.cfm?purl=10916490. ("interlibrary loan requests 'are to be filled in compliance with the U.S. Copyright Act and fair use provisions of the federal Copyright Act.'").

take away any rights existing under the fair use doctrine." H.R. Rep. No. 94-1476, at 74 (1976). The House Report's discussion of other parts of Section 108 reinforces the point that Section 108(f)(4)'s purpose was to prevent any implication that Section 108 limited fair use. In the context of Section 108(h), the House Report observed:

> Although subsection (h) generally removes musical, graphic, and audiovisual works from the specific exemptions of section 108, it is important to recognize that the doctrine of fair use under section 107 remains fully applicable to the photocopying or other reproduction of such works. H.R. Rep. No. 94-1476, at 78.

Copyright scholars agree that Section 108 does not limit the availability of Section 107 to libraries. 4 William Patry, *Patry on Copyright* § 11:3 (2011) ("[I]f for one reason or another, certain copying by a library does not qualify for the section 108 exemption …, the library's photocopying would be evaluated under the same criteria of section 107 as other asserted fair uses. This interpretation not only gives meaning to both sections but is fully in line with the earlier committee reports."); 4-13 Melville Nimmer & David Nimmer, *Nimmer on Copyright* § 13.05 (2011) ("[I]f a given library or archive does not qualify for the Section 108 exemption, or if a qualifying library or archive engages in photocopying practices that exceed the scope of the Section 108 exemption, the defense of fair use may still be available.").

Similarly, judicial opinions addressing the relationship between the Copyright Act's specific exceptions and fair use state that a defendant's failure to qualify for a specific exception does not prejudice its fair use rights. In *Sega Enters., Ltd. v. Accolade, Inc*., 977 F.2d 1510, 1520-21 (9th Cir. 1992), Sega argued that because Accolade's disassembly of Sega's computer program did not fall within the Section 117 exception relating to software, Accolade could not rely upon Section 107. Sega's position was that Section 117 "constitutes a legislative determination that any copying of a computer program *other* than that authorized by section 117 cannot be considered a fair use of that program under section 107." *Id.* The Ninth Circuit responded that this "argument verges on the frivolous. Each of the exclusive rights created by section 106 of the Copyright Act is expressly made subject to all of the limitations contained in sections 107 through 120." *Id.* at 1521. The court went on to observe that:

> sections 107 and 117 serve entirely different functions. Section 117 defines a narrow category of copying that is lawful *per se* …. The fact that Congress has not chosen to provide a *per se* exemption to section 106 for disassembly does not mean that particular instances of disassembly may not constitute fair use.

Id. Before the District Court, Appellants attempted to distinguish *Sega* on its facts, but the principle of specific exceptions not restricting fair use applies nonetheless. *See also Encyclopedia Britannica Educ. Corp. v. Crooks*, 447 F. Supp. 243, 249 n.7 (W.D.N.Y. 1978) ("The legislative history … makes clear that the statutory exemptions were intended to supplement rather than supersede the doctrine of fair use").

Contrary to the assertions of the Authors Guild and other rights holders, allowing fair use to supplement Section 108 does not read Section 108 out of the statute. Section 108 sets forth certain situations where a library can always make reproductions and distributions without the right holder's authorization. Some of these actions might be fair uses, but Section 108 provides legal certainty that encourages the library to proceed without conducting the more complex fair use analysis.[4] Other actions under Section 108 might be beyond what fair use would allow, yet Congress in its balancing of competing interests decided to permit them. Section 108(f)(4) clarifies that libraries can rely on Section 108 when they meet its detailed criteria and on Section 107 in other circumstances, when they satisfy its more general criteria.

D. Fair Use Sufficiently Updates Section 108.

Congress wrote Section 108 in the age of photocopiers. Recognizing that Section 108 might need to be updated, the Library of Congress in 2005 convened a group of publishers and librarians to examine Section 108 and make recommendations for how it should be amended to reflect the needs of libraries in the digital age. I was a member of

[4] *See also* Randolph D. Moss, Office of Legal Counsel, *Whether And Under What Circumstances Government Reproduction Of Copyrighted Materials Is A Noninfringing "Fair Use" Under Section 107 Of The Copyright Act Of 1976* 14 n.12 (1999). ("[S]ection 108 of the 1976 Act does not narrow the protection for fair use provided by the common law doctrine codified in section 107 …. Section 108 thus fairly can be viewed as a very valuable—and not superfluous—safe harbor: If a certain library practice is noninfringing under the specific and detailed provisions of section 108(a) … a library need not be concerned about how that particular photocopying practice would fare under section 107's more complex and indeterminate fair use standards.")

the Study Group. After three contentious years, the Section 108 Study Group issued a report that reflected at a high level agreement on some aspects of Section 108 that should be updated. The report, however, did not resolve many other important issues such as orphan works, mass digitization, and electronic reserves, nor did it propose statutory language in the areas where there was agreement.

Moreover, the Study Group did not conclude that fair use was inadequate to supplement Section 108's specific provisions. Indeed, the Study Group observed, "[i]n addition to section 108, libraries and archives rely upon fair use to make copies of copyrighted works for preservation and other purposes." *The Section 108 Study Group Report* 21 (2008). The Study Group stated that "section 108 was not intended to affect fair use. Certain preservation activities fall within the scope of fair use, regardless of whether they would be permitted by section 108." *Id.* at 22.

The difficulty of translating even the simplest of the Section 108 Study Group's recommendations into legislation was displayed at a symposium on Section 108 reform hosted by Columbia University Law School in February 2013.[5] For example, the suggestion that Section 108 be expanded to apply to museums engendered a debate concerning how museums should be defined, and the need to define libraries and archives, currently undefined in Section 108.[6]

Additionally, some of the Study Group's recommendations could have the effect of limiting what libraries do today. The Study Group, for instance, proposed a complex regulatory scheme for website archiving, an activity performed by libraries as well as commercial search engines. Indeed, at the Columbia symposium it was evident that some rights holders saw the "updating" of Section 108 as an opportunity to repeal Section 108(f)(4) and restrict the availability of fair use to libraries. This is completely unacceptable to libraries. In essence, these rights holders seek to deny libraries the benefit of the most significant privilege of the Copyright Act, which the Supreme Court recently described as part of "the traditional contours of copyright protection" and one of

[5] Symposium: *Copyright Exceptions for Libraries in the Digital Age: Section 108 Reform*, 36 COLUMBIA J. L. & ARTS 527 (2013).

[6] Even if consensus could be reached on modernizing Section 108, constantly evolving technology would quickly render it out of date once again.

copyright law's "built-in First Amendment accommodations." *Golan v. Holder*, 132 S. Ct. 873, 890 (2012).

The fact that Section 108 may reflect a pre-digital environment does not mean it is obsolete. It provides libraries and archives with important certainty with respect to the activities it covers. Furthermore, Section 108 provides courts with importance guidance concerning the application of Section 107. For example, in enacting Section 108(c), Congress indicated that there is a strong public policy interest in libraries making replacement copies. Accordingly, when a library makes a replacement copy that exceeds the specific provisions of Section 108(c) – for example, the library makes four copies rather than three copies – a court should give great weight to Congress's recognition of the public policy interest in replacement copies when assessing the first fair use factor: the purpose and character of the use.[7] To be sure, this "substantial compliance" with Section 108 is not outcome determinative. It simply tilts the first factor analysis in favor of the library.

III. LIBRARIES NO LONGER NEED ORPHAN WORKS LEGISLATION.

LCA has a long history of involvement in the orphan works issue. It provided extensive comments to the Copyright Office during the course of the Office's study that led to the Office's 2006 Orphan Works Report. LCA also actively participated in the negotiations concerning the orphan works legislation introduced in the 109th and the 110th Congresses. Although LCA strongly supported enactment of these bills, significant changes in the copyright landscape over the past eight years convince us that libraries no longer need legislative reform in order to make appropriate uses of orphan works.

A. The "Gatekeeper Problem" Has Diminished.

In its March 25, 2005, response to the Copyright Office's initial notice of inquiry concerning orphan works, LCA provided a long list of examples of the uses libraries sought to make of orphan works. It explained that while these uses "would significantly benefit the public without harming the copyright owner," copyright law nonetheless inhibited these uses. Even though it believed that many of these uses would qualify as fair use, "the uncertainty inherent in Section 107, when combined with the possibility of

[7] *See* Jonathan Band, *The Impact of Substantial Compliance with Copyright Exceptions on Fair Use*, 59 J. COPYRIGHT SOC'Y 453 (2012).

significant statutory damages notwithstanding the absence of actual damages, have caused various 'gatekeepers'—typically publishers or in-house counsel at universities—to forbid these uses." Since 2005, the "gatekeeper problem" has diminished markedly for the following reasons.

1. Fair use is less uncertain.

As previously noted, over the past eight years, courts have issued a series of expansive fair use decisions that have clarified its scope. In Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605 (2d Cir. 2006), Perfect 10 v. Amazon.com, 508 F.3d 1146 (9th Cir. 2007), and A.V. v. iParadigm, 562 F.3d 630, 639 (4th Cir. 2009), the courts found that the repurposing or recontextualizing of entire works by commercial entities was "transformative" within the meaning of fair use jurisprudence and therefore a fair use. Courts further recognized that a nonprofit educational purpose weighed heavily in favor of a fair use finding in Cambridge Univ. Press v. Becker, 863 F. Supp. 2d 1190 (N.D. Ga. 2012), Authors Guild, Inc. v. HathiTrust, No. 11 CV 6351, 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012), Ass'n for Info. Media and Equip. v. Regents of the Univ. of California, No. CV 10-9378 CBM (MANx), 2011 WL 7447148 (C.D.Cal. Oct. 3, 2011), and Ass'n for Info. Media and Equip. v. Regents of the Univ. of California, No. CV 10-9378 CBM (MANx) (C.D.Cal. Nov. 20, 2012). Relying on Perfect 10, iParadigm, and Bill Graham Archives, the general counsel of the U.S. Patent and Trademark Office (USPTO) opined that the copying of technical articles by the USPTO and patent applicants during the course of the patent examination process constituted fair use.[8] Importantly, Amazon.com, iParadigm, and HathiTrust all involved mass digitization.

All these uses were determined to constitute fair use even though the copyright owners were locatable. Gatekeepers at libraries and archives understand that similar uses of orphan works are all the more likely to fall within the fair use right because such uses would have no adverse effect on the potential market for the work.[9] Additionally, the

[8] Bernard Knight, USPTO General Counsel, *USPTO Position on Fair Use NPL Copies of Made in Patent Examination* (January 19, 2012) http://www.uspto.gov/about/offices/ogc/USPTOPositiononFairUse_of_CopiesofNPLMadeinPatentExamination.pdf.

[9] The second fair use factor, the nature of the copyrighted work, also weighs in favor of fair use when the work is an orphan. *See* Jennifer Urban, *How Fair Use Can Solve the Orphan Works*

Code of Best Practices in Fair Use for Academic and Research Libraries, developed by the Association of Research Libraries,[10] explicitly concludes that the orphan status of a work in a special collection enhances the likelihood that its use by a library is fair. The development of the *Code* was prompted by Professor Michael Madison's insight (following a review of numerous fair use decisions) that the courts were implicitly or explicitly, asking about habit, custom, and social context of the use, using what Madison termed a 'pattern-oriented' approach to fair use reasoning. If the use was normal in a community, and you could understand how it was different from the original market use, then judges typically decided for fair use.[11]

Based on this insight, the Association of Research Libraries undertook an effort to "document[] the considered views of the library community about best practices in fair use, drawn from the actual practices and experience of the library community itself."[12] The resulting *Code of Best Practices* identified "situations that represent the library community's current consensus about acceptable practices for the fair use of copyrighted materials and describes a carefully derived consensus within the library community about how those rights should apply in certain recurrent situations." Id.

One of the *Code's* principles directly addresses the digitizing and the making available of materials in a library's special collections and archives. The *Code* states that the fair use case for such uses "will be even stronger where items to be digitized consist largely of works, such as personal photographs, correspondence, or ephemera, whose owners are not exploiting the material commercially and likely could not be located to seek permission for new uses." Id. at 20. That is, the fair use case is stronger for orphan works. Significantly, the *Code* does not require a library to search for the copyright owner of such non-commercial material prior to digitizing it. Rather, the *Code* trusts librarians to exercise their professional judgment and expertise to determine whether the copyright owners of such materials are likely to be unlocateable, i.e., to presume

---

Problem, 27 Berkeley Technology Law Journal 1379 (2012), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2089526.

[10] The Code has been endorsed by the American Library Association, the Association of College and Research Libraries, the Arts Libraries Society of North America, the College Art Association, the Visual Resources Association, and the Music Library Association.

[11] Patricia Aufderheide and Peter Jaszi, Reclaiming Fair Use 71 (2011).

[12] Association of Research Libraries, et al., Code of Best Practices in Fair Use for Academic and Research Libraries 3 (2012).

responsibly that certain types of works are orphans.

2. Injunctions are less likely.

Historically, courts routinely issued injunctions when they found copyright infringement, presuming that the injury caused was irreparable. In 2006, however, the Supreme Court in eBay v. MercExchange, 547 U.S. 388 (2006), ruled that courts should not automatically issue injunctions in cases of patent infringement, but instead should consider the four factors traditionally employed to determine whether to enjoin conduct, including whether the injury was irreparable and whether money damages were inadequate to compensate for that injury. Lower courts in cases such as Salinger v. Colting, 607 F.3d 68 (2d Cir. 2010), have held that the Supreme Court's reasoning in eBay applies to the Copyright Act was well. The abolishment of the automatic injunction rule diminishes the probability that a court will enjoin a library's use of an orphan work in the unlikely event that the court finds the use to infringe; the copyright owner bears the heavy burden of proving that the library's use causes her irreparable injury.

3. Mass digitization is more common.

The leading search engines, operated by two of the world's most profitable companies, routinely cache billions of web pages without the copyright owners' permission.[13] This industry practice has faced absolutely no legal challenge in the United States since the Amazon.com decision in 2007, cited above. Gatekeepers understand that a court would favorably evaluate a non-profit library's fair use defense in the context of this industry practice.

Moreover, in part because of the legal developments described above, libraries across the country have begun engaging in the mass digitization of special collections and archives.[14] The more they engage in these activities, the more confident libraries—and their gatekeepers—become with their fair use analysis concerning the mass digitization of presumptively orphan works.

The controversy concerning the HathiTrust Orphan Works Project (OWP) has not shaken this confidence. In 2011, the University of Michigan (UM) announced an orphan works project, under which it would make orphaned books digitally available to

[13] http://www.worldwidewebsize.com/.
[14] The appendix contains a description of one such project by the New York Public Library.

authorized users of HathiTrust member libraries that had those books in their collections.[15] Several HathiTrust member libraries joined UM in this pilot project. The UM Library developed a procedure to identify books in copyright that were not on the market and for which a rights holder could not be identified or located. The procedure included the listing of possible orphan works on a website to provide copyright owners with the opportunity to claim the works. After UM posted a list of 150 possibly orphaned books, the Authors Guild re-posted the list to its blog, whose readers helped the Guild locate the authors of several of the books (but few copyright owners). Shortly thereafter, the Authors Guild initiated a copyright infringement action against UM, the HathiTrust, and some of the other libraries that participated in the orphan works pilot. In response, HathiTrust suspended the orphan works project.[16]

This high profile litigation concerning possibly orphaned books has not deterred libraries from engaging in the mass digitization of archives and special collections. The subject matter of these mass digitization projects is completely different from the published books at issue in the HathiTrust case. Much, if not all, of these historical records, photographs, and ephemera have never been distributed commercially. The HathiTrust litigation, thus, has helped delineate for libraries which orphan works projects will subject them to greater risk of infringement litigation. Moreover, the litigation has demonstrated the ultimate futility of the "reasonably diligent search" approach embodied by the orphan works legislation in the 109th and 110th Congresses. Using the crowd-sourcing power of the Internet and the publicity of the litigation, the Authors Guild was able to generate more information more quickly than a small team of individuals consulting existing databases and search engines. A copyright owner will always be able to identify a trail that would have led the user to his doorstep, and the user's only defense

[15] Critics of the OWP often mischaracterized the nature of the project by suggesting it would have made entire works downloadable by anyone on the open web. In reality, access to the text of orphan works under the OWP would have been limited to viewing or printing one page at a time on a web browser window while logged in and authenticated as a university library user—and even then the OWP would only allow as many simultaneous users as there were hard copies in the library's collection.

[16] The district court ultimately found that the infringement claim regarding the OWP was moot because the OWP had been suspended. Notwithstanding the suspension of the OWP, I continue to believe that it was a fair use. *See* Resource Packet on Orphan Works: Legal and Policy Issues for Research Libraries, http://www.arl.org/bm~doc/resource_orphanworks_13sept11.pdf.

would be that she did not have the resources to explore every fork that she would have encountered along the way.[17]

B. Profound Disagreement Remains.

In 2013, the Copyright Office issued a Notice of Inquiry Concerning Orphan Works and Mass Digitization. The significant diversity of opinion expressed in the comments submitted in the response to the Notice of Inquiry indicates that it will be extremely difficult to forge a consensus approach to these issues. Indeed, the comments are literally all over the map. There was less agreement in 2013 than in 2007 both on the existence of a problem and the best approach to solve it.

Last month, the Copyright Office held a public meeting on orphan works and mass digitization. The public meeting indicated that nothing has changed over the past year. The divisions are just as profound now as they were in the initial round of comments.[18] Indeed, the divisions between different communities may be even deeper now than before. The public meeting revealed fundamental disagreement as to whether the Constitutional rationale of the copyright system is to promote public benefit. Likewise, the meeting exposed a basic divergence concerning the correctness of fair use decisions over the past decade. In fact, one rights holder representative compared the recent fair use case law to *Plessy v. Ferguson*, suggesting that these fair use holdings were as legally and morally flawed as the Supreme Court's 1892 ruling upholding the "separate but equal" doctrine. The inflammatory nature of this analogy was exceeded only by another rights holder representative threatening three times during the course of one panel to sue libraries if they engaged in additional mass digitization activities. The hostility exhibited by some rights holders to users in general, and libraries in particular,

---

[17] *See, e.g.*, *Authors Guild v. Google Inc*., 770 F. Supp. 2d 666, 681 ("From Google's point of view, [my grandfather's memoir] is an 'orphaned' book" because the company "is likely to be unsuccessful in trying to locate the publisher, since the book was self-published and my grandfather is now deceased," but "[f]rom my family's point of view, [the memoir] is not orphaned at all. It is very clear who owns the copyright."). Additionally, libraries now have far more experience than in 2005 with searching for the copyright owners of material in archives and special collections. These searches are more time consuming, expensive, and inconclusive than we believed in 2005. This further reinforces the importance of trusting librarians' professional judgment (rather than item-by-item searching) to conduct fair use analysis for mass digitization projects.

[18] For a more detailed discussion of the different points of view expressed at the public meeting, *see* http://policynotes.arl.org/post/79876737815/recap-of-the-copyright-offices-roundtables-on-orphan

suggests that any legislative process concerning orphan works, mass digitization, or Section 108 is bound to fail.

IV. THE HATHITRUST CASE.

The *HathiTrust* case is one of several cases resulting from Google Book Search. The Authors Guild unfortunately sued a consortium of libraries for copyright infringement, but fortunately the district court found that the consortium's activities were permitted under the fair use doctrine and Section 121, an exception for the benefit of people with print disabilities.

Starting in 2004, Google entered into partnerships with leading research libraries, under which it borrowed millions of books from the libraries, scanned the books into its search database, and provided the libraries with digital copies of the books they had borrowed. Columbia was once of these libraries. The search results Google provided in response to a query would be a list of books that contained that search term. If a user clicked on a particular book, Google would display three "snippets" a few sentences long containing the search term, as well as bibliographic information concerning the book. In 2005, the Authors Guild (AG) and five publishers sued Google for infringing copyright by scanning the books into its search database. The parties began settlement negotiations, and reached a complex class action settlement agreement in 2008.[19] Among its many provisions, the settlement agreement allowed Google's partner libraries to create a "Research Corpus," a set of all the scans made by Google in connection with the Library Project. This Research Corpus makes up the core of the HathiTrust Digital Library (HDL). After a lengthy review process, the presiding judge, Denny Chin, rejected the settlement in 2011.[20]

---

[19] For a more detailed discussion of the settlement and the litigation leading up to it, see Jonathan Band, The Long and Winding Road to the Google Books Settlement, 9 J. MARSHALL REV. INTELL. PROP. L. 227 (2009).

[20] For a more detailed discussion of the Judge Chin's rejection of the settlement, see Jonathan Band, A Guide For the Perplexed Part IV:
The Rejection of the Google Books Settlement,
http://www.librarycopyrightalliance.org/bm%7Edoc/guideiv-final-1.pdf (March 2011).

Once the settlement was rejected, AG resumed its litigation against Google.[21] Additionally, AG and several authors' associations in other countries including Canada, Australia, and Norway separately sued HDL for copyright infringement.

HDL envisioned several uses of its database: preservation; full-text searches; and full-text access only for the print disabled. Moreover, as discussed above, HDL announced an orphan works pilot program. (OWP). AG claimed that the copies of the books HDL made when it created the database (i.e., when the digital files were transmitted by Google) infringed copyright. AG additionally claimed that the OWP would infringe copyright.

HDL promptly suspended the OWP. AG, however, continued to pursue the litigation. The central legal issue was whether the copies made by HDL were a fair use. In addition to arguing that these copies were not fair use, AG asserted that HDL could not even raise fair use as a defense because libraries could only engage in the copying permitted under 17 U.S.C. § 108, the specific exception for libraries.[22] For its part, HDL argued that the copies it made were justified by the purposes of its mass digitization project: preservation, search, and access for the print disabled. Because the OWP had been suspended indefinitely before any works had been made available, HDL argued the program was moot.

Ruling on a motion for summary judgment on October 12, 2012, Judge Baer of the U.S. District Court for the Southern District of New York decided in favor of HDL on virtually every issue. He found that the library specific exceptions in section 108 do not restrict the availability to libraries of fair use under section 107.

With respect to fair use itself, Judge Baer declared: "I cannot imagine a definition of fair use that would not encompass the transformative uses made by [HDL] and would require that I terminate this invaluable contribution to the progress of science and cultivation of the arts that at the same time effectuates the ideals espoused by the [Americans with Disabilities Act]."[23] In the course of reaching this fair use conclusion, Judge Baer made the following findings:

---

[21] The publishers reached a narrow settlement agreement with Google that did not require judicial approval.
[22] This issue is discussed in greater detail above in Section II.
[23] Authors Guild, Inc. v. HathiTrust, 902 F.Supp.2d 445, 464 (S.D.N.Y. 1992)

- The creation of a search index is a transformative (and therefore favored) use under the first fair use factor: "The use to which the works in HDL are put is transformative because the copies serve an entirely different purpose than the original works: the purpose is superior search capabilities rather than actual access to copyrighted material."[24]
- The use of digital copies to facilitate access for the print-disabled is also transformative. Because print-disabled persons are not a significant potential market for publishers, providing them with access is not the intended use of the original work.
- HDL enabled libraries to "preserve their collections in the face of normal deterioration during circulation, natural disasters, or other catastrophes that decimate library collections, as well as loss due to theft or misplacement." Judge Baer quoted the House Judiciary Committee report on the 1976 Copyright Act stating that the efforts of libraries and archives "to rescue and preserve this irreplaceable contribution to our cultural life are to be applauded, and the making of duplicate copies for purposes of archival preservation certainly falls within the scope of 'fair use.'"
- The AG failed to show that HDL created any security risks that threatened AG's market.
- AG's suggestion that HDL undermines existing and emerging licensing opportunities is "conjecture."
- The goal of copyright to promote the progress of science are better served by allowing HDL's use than by preventing it.

In addition, Judge Baer found that the specific exception for the print disabled, the Chafee Amendment, 17 U.S.C. § 121, allowed HDL to provide full text access to the print disabled. The court found that HDL was an "authorized entity" within the meaning of the statute because it had a primary mission of providing services to the print disabled. Moreover, the digital copies met the definition of "specialized formats" because they were made available only to the print disabled.

In sum, the court found two means of getting to the same objective: providing

[24] Id. at 460.

accessible copies to the print disabled. On the one hand, the mass digitization of over 10 million published books, of which at least 70 percent were still in copyright, for the purpose of providing accessible format copies to the print disabled was a fair use. Notwithstanding the large number of works, the judge didn't see this as a hard case; the fairness of the use was obvious. On the other hand, the Chafee Amendment also permitted libraries to make accessible format copies for print disabled students and faculty.

AG has appealed Judge Baer's decision to the U.S. Court of Appeals for the Second Circuit. The oral argument was held on October 30, 2013. I hope that the Second Circuit will agree with Judge Baer that HDL preserves important works, allows them to be searched, and provides access to the print disabled, without causing any economic harm to rights holders.

V. LEGISLATIVE RECOMMENDATIONS.

Although we do not seek amendment of Section 108 or special legislation targeting orphan works, certain narrow changes to the Copyright Act would benefit libraries and other cultural heritage institutions. For example, Section 504(c)(2) allows for the remission of statutory damages to libraries, educational institutions, and public broadcasters when they reasonably believed that certain activities were fair uses. However, this limitation does not apply to museums, and it should. Moreover, the limitation for libraries and educational institutions applies only to infringements of the reproduction right, not the performance, display, distribution, or derivative work right. As a result, the limitation provides little benefit, particularly for Internet uses that involve the display of a work on a website. The remission provision for non-profit institutions should apply to museums and to infringements of all exclusive rights under Section 106.

LCA will address other potential amendments as the subcommittee considers the relevant sections, e.g., the anti-circumvention provisions of the Digital Millennium Copyright Act, 17 U.S.C. § 1201. However, two issues should receive special mention because of their significance to libraries. First, the preemption of contractual provisions limiting copyright exceptions. An increasing proportion of library acquisitions is digital resources. Indeed, many research libraries spend over 65% of their acquisition budgets on electronic resources. These licenses often contain terms that restrict fair use, first sale,

and other user rights under the Copyright Act. As it reviews the Copyright Act, the subcommittee should consider possible amendments to Section 301(a) to ensure that libraries and other cultural heritage institutions will be able to preserve digital materials in their collections, notwithstanding contractual provisions to the contrary.

Second, the subcommittee should consider the impact of the Copyright Act on people with disabilities. As previously mentioned, Judge Baer in the HathiTrust case found that both fair use and Section 121 permitted HathiTrust to provide print disabled faculty with students with access to the full text of books within the HathiTrust database. The subcommittee should consider whether Section 121 should be expanded to apply to people with other disabilities.

April 2, 2014

APPENDIX

The New York World's Fair of 1939 and 1940 (the "Fair") was held in Flushing Meadows in Queens. At the conclusion of the Fair, the corporation in charge of the Fair dissolved and donated a large amount of material to the New York Public Library ("NYPL"). The corporation donated over 2,500 boxes of records and documents, as well 12,000 promotional photographs. These records document an important event in history and are heavily used by researchers and the public.

When deciding whether to digitize this collection and make it available online, NYPL conducted a thorough, good-faith search for rights holders. It started by trying to determine the copyright status for the nearly ten tons of works in the collection. The publication status of much of the material was difficult to determine and was, therefore, treated as if it were in copyright. Because the material may be in copyright, NYPL shifted its focus to find a copyright owner. It spent days combing through the legal records of the Fair to determine whether the Fair's copyright was ever assigned to a third party. It also tried to determine whether copyrights were assigned at the dissolution of the corporation, but could not find an answer in the archive. When the records of the Fair did not help, NYPL searched for rights holders utilizing other methods, including searches on Google, the Copyright Office records, and other relevant sources. This search was time-consuming and, ultimately, fruitless.

NYPL could not locate a rights holder who owned the rights to the material in the collection. After balancing the educational benefit of digitizing and making portions of the collection available online with the risk that a rights holder might subsequently surface, NYPL determined to move forward with the project, guided by fair use considerations. The potential maximum copyright liability for this project was estimated to be in excess $1.8 billion dollars. Despite this potential liability, NYPL not only digitized and posted the collection, it used the material in a free app that was later named one of Apple's "Top Education Apps" of 2011. Furthermore, an educational curriculum has been built around this material.

So far, no rights holder has contacted NYPL to ask that it limit the uses of works from the Fair collection. If a rights holder wished to contact NYPL about its uses, NYPL has made its contact information available online and in the iPad application.