

Research Libraries and New Technologies, Promoting Access to Information, Learning, and Innovation for Today and the Future

KRISTA L. COX

CONTENTS

I.	BACKGROUND: ROLE OF RESEARCH LIBRARIES HISTORICALLY AND TODAY.....	261
II.	FAIR USE PROTECTS ABILITY OF LIBRARIES TO ADAPT TO NEW TECHNOLOGIES TO PROVIDE ACCESS TO INFORMATION.....	263
	A. <i>Fair Use/Transformative</i>	264
	B. <i>Text and Data Mining</i>	266
III.	ARE EXTENDED COLLECTIVE LICENSING REGIMES NECESSARY?	269
IV.	ARE FAIR USE SAVINGS CLAUSES ENOUGH?.....	277
V.	LIBRARY ACTIVITY TO PROMOTE ACCESS TODAY	281
	A. <i>Digitization of Special Collections</i>	281
	B. <i>Beyond Print</i>	285
	C. <i>Accessibility</i>	285
	D. <i>Preservation of Websites and 21st Century Ephemera</i>	288
	E. <i>Institutional Repositories</i>	291
VI.	CONCLUSION	294

I. BACKGROUND: ROLE OF RESEARCH LIBRARIES HISTORICALLY AND TODAY

The role of libraries today is rooted in their historical mission. At their heart, libraries provide access to knowledge and information. They do so by preserving the cultural and historical record, not only for today, but for generations to come. They support teaching, learning and research by providing access to works, curating collections and ensuring that connections between different materials

can be made. Libraries also have a long history of providing physical spaces for studying, meeting or for exhibitions. They have also provided accessible formats to individuals with print and other disabilities, helping to ensure that everyone can access information. While these activities represent the historical role for libraries, these same endeavors hold true today.

Beyond this historic role, libraries today have also evolved to accommodate the new ways of learning that technology has permitted. Learning is often more interdisciplinary today with the advent of text and data mining which allow new connections to be made across disciplines. Libraries continue to provide access to new technologies including computers and the Internet as the digital age came into being,¹ and maker spaces with 3-D printers today.² As such, they are hubs of creation and exploration, allowing individuals to use technologies that they would not otherwise be able to access. While libraries have always understood the importance of preservation, this role has an increasingly important role with the growth of websites, social media and other ephemera produced today.³ Additionally, with the rapid evolution of technology, the risk of obsolescence of particular formats arises.

Ultimately, libraries have adapted to the changing landscape to ensure that they fulfill their missions of preservation and provision of access to knowledge and culture. Libraries lie at the heart of discovery

¹ See, e.g., AM. LIBR. ASS'N, INTERNET CONNECTIVITY IN U.S. PUBLIC LIBRARIES (2009), http://www.ala.org/research/sites/ala.org.research/files/content/initiatives/plftas/issues/briefs/connectivitybrief_2009_10_final.pdf [<https://perma.cc/9GQH-R9AX>].

² Joseph Leahy, *Libraries Make Space for 3-D Printers; Rules are Sure to Follow*, NPR (Apr. 29, 2015), www.npr.org/sections/alltechconsidered/2015/04/29/401236656/libraries-make-space-for-3-d-printers-rules-are-sure-to-follow [<https://perma.cc/UAH6-4ZDC>].

³ See, e.g., *Hearing of Preservation and Reuse of Copyrighted Works Before the Subcomm. on Courts, Intellectual Property & the Internet of the H. Comm. on the Judiciary*, 113th Cong. (2014), <https://judiciary.house.gov/wp-content/uploads/2016/02/Neal-Columbia-Preservation-Reuse-Testimony-1.pdf> [<https://perma.cc/K4ZX-7GD3>] (Statement of James G. Neal, Vice President for Information Services and University Librarian of Columbia University in the City of New York) (“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.”).

and learning, promoting scholarly breakthrough by ensuring that individuals have access to the resources needed.

II. FAIR USE PROTECTS ABILITY OF LIBRARIES TO ADAPT TO NEW TECHNOLOGIES TO PROVIDE ACCESS TO INFORMATION

Fair use has accommodated new technologies and the ways in which individuals access information today. This important doctrine has allowed, for example, for the growth of searchable databases, text and data mining, and greater availability of accessible formats for those with print disabilities and more.

Unlike most of the Copyright Act, fair use is a flexible doctrine and is thus adaptable to the evolution of technology. Because the current Copyright Act was enacted in 1976, with some updates in 1998 for the digital age, some of its provisions may seem outdated as they were largely designed for an analog world. However, the law was written to be technology neutral. Additionally, fair use provides the mechanism to ensure that constant revisions to the Copyright Act are unnecessary. Fair use does not provide an exhaustive list of exceptions, but rather uses four factors to determine whether a use is permissible.

The flexibility of the fair use doctrine has been essential to the growth of the technology industry⁴ and, as a result of advancements in technology, to broader dissemination of information and new learning systems. It has enabled new advancements, such as search engines,⁵ permitted creativity,⁶ and allowed for scholarship.⁷ The use of search engines and ability to text and data mine are just two examples of fair uses that are available today, but would not have been conceived of when fair use was codified in statute in 1976. Allowing copyright law

⁴ See CCIA, FAIR USE IN THE US ECONOMY, <http://www.cciagnet.org/wp-content/uploads/library/CCIA-FairUseintheUSEconomy-2011.pdf> [<https://perma.cc/KP29-NDSX>] (Experts estimated that industries reliant on fair use contributed \$2.4 trillion to the United States economy in 2008-2009, representing approximately 17 percent of the United States GDP).

⁵ See 17 U.S.C. § 107 (2016); see also *Kelly v. Arriba Soft*, 336 F.3d 811 (9th Cir. 2003); *Perfect 10 v. Amazon*, 508 F.3d 1146 (9th Cir. 2007) (“electronic reference tool” is a highly transformative fair use).

⁶ See, e.g., *Campbell v. Acuff-Rose Music*, 510 U.S. 569 (1994).

⁷ The chapeau to Section 107 of the Copyright Act specifically points to examples such as criticism, comment, teaching, scholarship, and research as likely to support fair use.

to adapt to these technologies through fair use ensures that the law does not remain stagnant and allows creativity to flourish.

A. *Fair Use/Transformative*

The fair use doctrine provides an express limitation or exception to copyright, allowing the use of a copyrighted work without the rightholder's consent under certain circumstances. It is an essential element to copyright law in the United States, allowing the law to adapt to new technologies⁸ and accommodate freedom of speech.⁹ It serves as a vital balancing feature that supports the very purpose of the intellectual property system as laid out in the U.S. Constitution: "To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."¹⁰ Ultimately, the purpose of copyright is to benefit the public.¹¹

In order to ensure that the public interest is protected and that a rightholder's monopoly does not unnecessarily hinder progress and creativity, application of fair use has long been part of the United States copyright system.¹² While fair use has long been a part of

⁸ See, e.g., *Sony Corp. of America v. Universal City Studios*, 464 U.S. 417 (1984).

⁹ See, e.g., *Harper & Row Publishers v. Nation Enterprises*, 471 U.S. 539 (1985) ("the Framers intended copyright itself to be the engine of free expression.").

¹⁰ U.S. CONST. art. 1, § 8, cl. 8.

¹¹ The Supreme Court of the United States has confirmed that the larger goal of copyright is to benefit the public. See *Harper & Row*, 471 U.S. at 546 ("The monopoly created by copyright thus reward the individual author in order to benefit the public."); *Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975) ("The limited scope of the copyright holder's statutory monopoly . . . reflects a balance of competing claims upon the public interest: Creative work is to be encouraged and rewarded, but private motivation must ultimately serve the cause of promoting broad public availability of literature, music and other arts. The immediate effect of our copyright law is to secure a fair return for an author's creative labor. But the ultimate aim is, by this incentive, to stimulate artistic creativity for the general public good.").

¹² Justice Story articulated the basis for determining whether a use would be considered fair use noting that, "In short, we must often . . . look to the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects, of the original work." *Folsom v. Marsh*, 9 F. Cas. 342, 348 (C.C.D. Mass. 1841).

copyright jurisprudence, the statutory fair use factors were not codified until the 1976 Copyright Act.¹³

Fair use is essential to the United States' copyright system. It acts as a "safety valve" of copyright law, which "permits [and requires] courts to avoid rigid application of the copyright statute when on occasion, it would stifle the very creativity which that law is designed to foster."¹⁴ It also serves as a "built-in First Amendment accommodation[] . . . [and] allows the public to use not only facts and ideas contained in a copyright work, but also expression itself in certain circumstances."¹⁵

Whether a particular use is "transformative" is an important consideration. As the Supreme Court has noted, permitting transformative works furthers the Constitutional purpose of the intellectual property system:

The goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works. Such works thus lie at the heart of the fair use doctrine's guarantee of breathing space within the confines of copyright, and the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.¹⁶

The Supreme Court has determined that a transformative use is one which "adds something new, with a further purpose or different character, altering the first with new expression meaning or message."¹⁷ Courts have noted that for a use to be transformative, it must do "more than repackage or republish the original copyrighted work."¹⁸ Such a use must not substitute for the original work, but

¹³ 17 U.S.C. § 107 (2016).

¹⁴ *Campbell v. Acuff-Rose Music*, 510 U.S. 569, 577 (1994).

¹⁵ *Eldred v. Ashcroft*, 537 U.S. 186, 219-21 (2003).

¹⁶ *Campbell*, 510 U.S. at 579 (internal citations omitted).

¹⁷ *Id.* at 579.

¹⁸ *Authors Guild v. HathiTrust*, 755 F.3d 87, 96 (2d Cir. 2014).

instead serve a new, different function¹⁹ or expand its utility.²⁰ A key question is, therefore, whether the original work is being used to serve a new function; it is not necessary to make changes to the work itself to be considered transformative.²¹

B. *Text and Data Mining*

The advent of computers and searchable databases has allowed for greater computational analysis known as text and data mining. Researchers can detect patterns in results through these large-scale analyses that would not be possible without computer programs to crawl millions of articles or other digital content. In doing so, researchers can analyze vast amounts of information, identify patterns and trends, and understand connections between different texts that would not be possible without the use of text and data mining.

In order to facilitate text and data mining, large databases must be created which includes the digitization of materials that is not “born digital” to allow computers to read this information. Fair use jurisprudence has become considerably more certain in recent years, permitting the creation of searchable databases and text and data mining. In the recent case of *Authors Guild v. HathiTrust*, the Second Circuit called the creation of the full-text searchable database a “quintessentially transformative use.”²² HathiTrust involves a partnership with what is now over 110 research libraries across the world to aggregate digitized books and, under the program challenged by the Authors Guild, allowed users to search the aggregated database for specific terms.

The Second Circuit again concluded that Google’s creation of a searchable database—as well as the display of “snippets”—was fair use in its 2015 opinion in *Authors Guild v. Google*.²³ Yet even prior to

¹⁹ *Id.* at 96.

²⁰ *Authors Guild v. Google*, 804 F.3d 202 (2d Cir. 2015).

²¹ See *A.V. ex. rel. Vanderhuy v. iParadigms, LLC*, 562 F.3d 630, 639-40 (4th Cir. 2009) (“The question of whether a use is transformative does not rise or fall on whether the use perfectly achieves its intended purpose. The use of a copyrighted work need not alter or augment the work to be transformative in nature. Rather, it can be transformative in function or purpose without altering or actually adding to the original work.”).

²² *HathiTrust*, 755 F.3d 87.

²³ *Google*, 804 F.3d 202.

these recent cases, numerous courts have set strong precedent that searchable databases and text and data mining are permitted as fair uses under copyright law.²⁴ These cases have involved a variety of purposes, ranging from research by scholars, legal research, use by politicians and anti-plagiarism software, among others.²⁵ Given the variety of purposes that have been upheld and the quintessentially transformative nature of searchable databases, the creation of such databases is a fair use.

In the Second Circuit's recent decisions in the Authors Guild's litigation, the court repeatedly explains the importance of transformativeness in determining whether a use is fair. While transformativeness is not dispositive in determining whether the first factor favors fair use, "transformative uses tend to favor a fair use finding because a transformative use is one that communicates something new and different from the original or expands its utility, thus serving copyright's overall objective of contributing to public knowledge."²⁶ In the case of a searchable database, the transformative use "provides otherwise unavailable information about the originals."²⁷

While searchable databases provide useful information, "added value or utility is not the test: a transformative work is one that serves a new and different function from the original and is not a substitute

²⁴ See, e.g., *Kelly v. Arriba Soft*, 336 F.3d 811 (9th Cir. 2003) (holding that thumbnails of and in-line linking to images hosted on a photographers website in a search engine was fair use); *Perfect 10 v. Amazon*, 508 F.3d 1146 (9th Cir. 2007) (holding that Google's use of thumbnail versions of copyrighted images and in-line linking to the full images is an "electronic reference too" that constitutes a highly transformative fair use); see *A.V. ex. rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630 (creation of a database allowing teachers to compare a student's work with content available on the Internet to determine whether the work was plagiarized is fair use); *Field v. Google*, 412 F. Supp. 2d 1106 (D. NV. 2006) (copies of content in its website cache, including archival copies, used for web comparisons or identification in a search query is fair use); *White v. West*, No. 12 Civ. 1340 (JSR), WL 3385480 (S.D.N.Y. 2014) (copying of legal filings into search databases and the addition of metadata creates an interactive legal research tool that is a transformative fair use); *Fox v. TVEyes*, 43 F. Supp. 3d 379 (S.D.N.Y. 2014) (recording of entire contents of television and radio broadcasts to create a searchable database of the content is a fair use).

²⁵ See KRISTA L. COX, ASS'N OF RES. LIBR., ISSUE BRIEF: TEXT AND DATA MINING AND FAIR USE IN THE UNITED STATES (2015), <http://www.arl.org/storage/documents/TDM-5JUNE2015.pdf> [<https://perma.cc/BCL3-SJCT>].

²⁶ *Google*, 804 F.3d at 214.

²⁷ *Id.* at 215.

for it . . . the result of a word search is different in purpose, character, expression, meaning, and message from the page (and the book) from which it is drawn.”²⁸ Thus, the ingestion of works into a database creates a completely different function from the original work and is therefore transformative.

With respect to the second factor, the Second Circuit has noted that it “has rarely played a significant role in the determination of a fair use dispute.”²⁹ Indeed, courts have rarely found the second factor to weigh heavily in the fair use analysis.³⁰

In considering the third fair use factor: the amount copied, text and data mining relies on ingesting the entire work into the database. The Second Circuit noted in *HathiTrust* that

it was reasonably necessary for the [HathiTrust Digital Library] to make use of the entirety of the works in order to enable the full-text search function” and “we do not believe the copying was excessive . . . these copies are . . . reasonably necessary in order to facilitate . . . legitimate uses.³¹

Similarly, in *Authors Guild v. Google*, the Second Circuit confirmed:

As with *HathiTrust*, not only is the copying of the totality of the original reasonably appropriate to Google’s transformative purpose, it is literally necessary to achieve that purpose. If Google copied less than the totality of the originals, its search function could not advise searchers reliably whether the searched term appears in a book (or how many times).³²

²⁸ *Authors Guild v. HathiTrust*, 755 F.3d 87, 96-97 (2d Cir. 2014).

²⁹ *Google*, 804 F.3d at 220 (citing William F. Patry, PATRY ON FAIR USE § 4.1 (2015)).

³⁰ See, e.g., *On Davis v. The Gap*, 246 F.3d 152, 175 (2d Cir. 2001) (“The second statutory factor, the nature of the copyrighted work, is rarely found to be determinative.”); *Cambridge Univ. v. Patton*, 769 F.3d 1232, 1270 (11th Cir. 2014) (“[T]he second fair use factor is of relatively little importance in this case.”).

³¹ *HathiTrust*, 755 F.3d at 98.

³² *Google*, 804 F.3d at 221.

Courts have thus found that because verbatim copying of the entire work is necessary for searchable databases, this copying is reasonable and the third factor is neutral.

Finally, with respect to the fourth factor, courts have found that the highly transformative nature of the creation of a searchable database, and use in text and data mining, is not likely to have an adverse impact on the market of the original work because the use does not supersede the copyrighted work. The fourth factor is closely linked to the first factor and the more transformative the use, the less likely it is to act as a substitute for the original work.³³ As the Second Circuit noted in *HathiTrust*, searching the text of a book to determine whether key words are used within that text “does not serve as a substitute for the books that are being searched.”³⁴

In upholding the *HathiTrust* and *Google Books* projects, the Second Circuit ensures that copyright law has adapted to technology and truly promotes the progress of science and useful arts. Recent copyright law, including the Second Circuit decisions, has allowed libraries to fulfill their missions of collecting, organizing, preserving, and providing access to knowledge.

III. ARE EXTENDED COLLECTIVE LICENSING REGIMES NECESSARY?

The Copyright Office has been engaged in studies regarding orphan works—works for which it is difficult or impossible to locate the rightholder—and mass digitization, among other issues. Libraries participating in roundtables³⁵ and submitting written comments³⁶ on

³³ See *Campbell v. Acuff-Rose Music*, 510 U.S. 569, 591 (1994).

³⁴ *HathiTrust*, 755 F.3d at 100.

³⁵ *Orphan Works March 2014 Public Roundtables*, U.S. COPYRIGHT OFF., <http://www.copyright.gov/orphan/2014roundtable.html> [<https://perma.cc/42PR>].

³⁶ See *Comments on Orphan Works*, U.S. COPYRIGHT OFF. (2013), http://www.copyright.gov/orphan/comments/noi_10222012/ [<https://perma.cc/N5A4-7J5F>]; *Reply Comments on Orphan Works*, U.S. COPYRIGHT OFF. (2013), http://www.copyright.gov/orphan/comments/noi_11302012/ [<https://perma.cc/J34J-EX2N>]; *Additional Comments on Orphan Works*, U.S. COPYRIGHT OFF. (2014), http://www.copyright.gov/orphan/comments/Docket2012_12/ [<https://perma.cc/8B2G-BTAM>]. In particular, see *Libr. Copyright All., Comments of the Library Copyright Alliance in Response to the Copyright Office’s Notice of Inquiry Concerning Orphan Works and Mass Digitization* (May 16, 2011), http://www.copyright.gov/orphan/comments/noi_10222012/Library-Copyright-Alliance.pdf [<https://perma.cc/S32D-C4ZF>] (“Moreover, any legislative approach that

these issues cautioned against legislating in this area. For example, Duke University Libraries stated, “The one suggestion that certainly would not improve the current situation would be a collective licensing regime.”³⁷ Similarly, the University of Minnesota Libraries wrote, “we strongly oppose the idea of extended collective licensing as a solution to the problem of orphan works. Paying licensing fees in

involves licensing, such as extended collective licensing, is completely unacceptable to the library community. It would be enormously costly to users, and little if any of the fees collected would ever actually reach the copyright owners of the orphan works. Instead, fees would be consumed by the collecting societies’ administrative expenses and the cost of searching for absent owners”); Emory U. Libr., Comments of Emory University Libraries in Response to the Copyright Office’s Notice of Inquiry Concerning Orphan Works and Mass Digitization (Feb. 4, 2013),

http://www.copyright.gov/orphan/comments/noi_10222012/Emory_University_Libraries.pdf [<https://perma.cc/G37Z-GR5Q>]; Mass. Inst. of Tech. Libr., Additional Comments of Massachusetts Institute of Technology Libraries in Response to Orphan Works and Mass Digitization,

[http://www.copyright.gov/orphan/comments/Docket2012_12/Massachusetts-Institute-of-Technology\(MIT\)-Libraries.pdf](http://www.copyright.gov/orphan/comments/Docket2012_12/Massachusetts-Institute-of-Technology(MIT)-Libraries.pdf) [<https://perma.cc/BDA5-WB77>] (“We also do not believe collective agencies would offer an effective solution to managing orphan works. Because there is no copyright holder to pay if the work is truly an orphan, funds would not be fairly allocated. Indeed, the funds put forward would consist, essentially, of a “tax on socially beneficial uses,” as Melissa Levine, copyright officer at the University of Michigan, has commented. In addition, experience with collective agencies in Europe suggests that such a body may not be motivated to try sufficiently hard to find the copyright holder—so establishing this kind of model is likely to simply create a new stakeholder to monitor. In short, collective rights agencies are in our view not likely to result in the stated goal of direct compensation for copyright holders, as Jonathan Band demonstrated in his recent examination of licensing societies.”); N.Y. Pub. Libr., Reply Comments in Response to Orphan Works and Mass Digitization: Notice of Inquiry (Mar. 6, 2013),

http://www.copyright.gov/orphan/comments/noi_11302012/New-York-Public-Library.pdf [<https://perma.cc/Y5MD-7DSK>] (Noting that any potential legislation should “Not include extended collective licensing or other licensing schemes.”); UCLA Libr., Additional Comments of University of California, Los Angeles Libraries in Response to Orphan Works and Mass Digitization: Request for Additional Comments (May 21, 2014), [http://www.copyright.gov/orphan/comments/Docket2012_12/University-of-California-Los-Angeles\(UCLA\)-Library.pdf](http://www.copyright.gov/orphan/comments/Docket2012_12/University-of-California-Los-Angeles(UCLA)-Library.pdf) [<https://perma.cc/PCM3-XDBA>] (“In short, compulsory licensing is not the solution to the problem of orphan works. Instead, a combination of community best practices and reliance on the existing legal framework best furthers the mission of higher education and the promotion of science of useful arts.”); U. Minn. Libr., Comments of University of Minnesota Libraries (May 20, 2014),

http://www.copyright.gov/orphan/comments/Docket2012_12/University-of-Minnesota-Libraries.pdf [<https://perma.cc/6EU4-DK5A>].

³⁷ Duke U. Libr., Comments from the Duke University Libraries in response to the Copyright Office’s Notice of Inquiry concerning orphan works and mass digitization (Jan. 2013), https://www.copyright.gov/orphan/comments/noi_10222012/Duke-University-Libraries.pdf [<https://perma.cc/3W4F-EP27>].

escrow for rightsholders who do not exist or will never receive them is an exercise in waste. Moreover, due to the limited budgets of libraries, extended collective licensing would create a concrete barrier to digitization in the service of the public interest.”³⁸ Fair use jurisprudence in recent years has moved forward, diminishing the need for orphan works legislation. Furthermore, efforts to enact orphan works legislation in 2008 resulted in more complicated and burdensome proposals before these efforts ultimately failed.³⁹ With respect to mass digitization and extended collective licensing regimes, participants at the roundtables seemed almost universally opposed to extended collective licensing, noting the poor management of collecting societies and difficulties in tracking down rightsholders.⁴⁰

In June 2015, the Copyright Office released its report on Orphan Works and Mass Digitization, proposing legislation to address orphan works as well as a mass digitization/extended collective licensing pilot program.⁴¹ In its report, the Copyright Office rejects the idea that fair use can provide an adequate solution to the orphan works solution,⁴²

³⁸ U. Minn. Libr., *supra* note 36.

³⁹ H.R. 5889, 110th Cong. (2008).

⁴⁰ See Krista L. Cox, *Recap of the Copyright Office's Roundtables on Orphan Works and Mass Digitization*, ARL POL'Y NOTES (Mar. 17, 2014), <http://policynotes.arl.org/?p=51> [<https://perma.cc/A3GK-FNC5>]; LIBR. COPYRIGHT ALL., ADDITIONAL COMMENTS OF THE LIBRARY COPYRIGHT ALLIANCE TO THE COPYRIGHT OFFICE'S NOTICE OF INQUIRY CONCERNING ORPHAN WORKS AND MASS DIGITIZATION (May 16, 2014), <http://www.librarycopyrightalliance.org/storage/documents/lca-additional-ow-comments-16may2014.pdf> [<https://perma.cc/LK5P-HCLL>] (“In contrast to the disagreement concerning orphan works legislation, there was general agreement at the public meeting that extended collective licensing (ECL) would not be an effective solution to issues relating to mass digitization, even if limited only to books.”).

⁴¹ U.S. COPYRIGHT OFF., ORPHAN WORKS AND MASS DIGITIZATION: A REPORT OF THE REGISTER OF COPYRIGHTS (2015), <http://copyright.gov/orphan/reports/orphan-works2015.pdf> [<https://perma.cc/W6CW-2LYD>].

⁴² *Id.* at 42-43 (“The judiciary has yet to explicitly address how to apply fair use to orphan works. Thus, the informed and scholarly views of some commenters as to the application of fair use in specific orphan works situations do not yet have as their basis any controlling case law. Also, fair use jurisprudence is, because of its flexibility and fact-specific nature, a less concrete foundation for the beneficial use of orphan works than legislation, and is always subject to change . . . the Office does not believe that reliance on judicial trends, which may turn at any point, is a sufficient basis to forgo a permanent legislative solution.”). *But see* Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2541 (2009) (“Fair use law is both more coherent and more predictable than many

and proposes a discussion draft for legislation largely based off the Shawn Bentley Orphan Works Act of 2008.⁴³ The discussion draft raises a number of concerns because of its burdensome nature and failure to appreciate the library community's comfort in relying on fair use.⁴⁴ In recent years, fair use has been clarified by courts to support broader application of fair use.⁴⁵ In addition to the further development of fair use jurisprudence, other areas of law, such as the elimination of automatic injunctions for infringement, have been clarified in favor of the users.⁴⁶ This recent jurisprudence has provided greater comfort to those in the library community who rely on fair use. Furthermore, with the development of best practices, such as the Code of Best Practices in Fair Use for Academic and Research Libraries⁴⁷ (among other relevant best practices in fair use),⁴⁸ users have greater

commentators have perceived once one recognizes that fair use cases fall into common patterns.”).

⁴³ The Shawn Bentley Orphan Works Act of 2008 was passed by the Senate, but failed in the House of Representatives. The discussion draft would provide a limitation on remedies, require a good faith diligent search, provides a notice of use requirement, and includes a fair use savings clause.

⁴⁴ Krista L. Cox, *Copyright Office Releases Report on Orphan Works and Mass Digitization; Recommends Burdensome Legislation*, ARL POL'Y NOTES, <http://policynotes.arl.org/?p=1075> [<https://perma.cc/65DR-ZTJX>] (“[T]he draft legislation has significant problems including overly burdensome and complicated requirements. The requirements for a reasonably diligent search and limitations on injunctions are highly problematic. Finally, the notice of use provision is as poisonous now as it was in 2008. If it is included in an orphan works provision, it will ensure that the provision is rarely, if ever, used.”).

⁴⁵ See *supra* Sections II.A, II.B.

⁴⁶ *eBay v. MercExchange*, 547 U.S. 388 (2006). In addition to the recent fair use jurisprudence upholding several transformative uses, the Supreme Court's decision in *eBay v. MercExchange* that patent right holders were not automatically entitled to an injunction in infringement cases, a decision which lower courts have extended to the copyright context, provides further comfort in relying on orphan works.

⁴⁷ ASS'N OF RES. LIBR., *CODE OF BEST PRACTICES IN FAIR USE FOR ACADEMIC AND RESEARCH LIBRARIES* (2012), <http://www.arl.org/storage/documents/publications/code-of-best-practices-fair-use.pdf> [<https://perma.cc/U2P9-4KUB>].

⁴⁸ See *Codes of Best Practices*, CTR. FOR MEDIA AND SOCIAL IMPACT, <http://cmsimpact.org/codes-of-best-practices/> [<https://perma.cc/4QSM-BNCB>] (including Set of Principles in Fair Use for Journalism; Code of Best Practices in Fair Use for Online Video; Best Practices in Fair Use In Teaching for Film and Media Educators; Fair Use of Images for Teaching, Research and Study; Code of Best Practices in Fair Use for Scholarly Research in Communication; Best Practices in Fair Use of Collections Containing

guidance as to what constitutes fair use. The Code of Best Practices for Academic and Research Libraries, for example, lays out eight statements of high-level principles of fair uses, with limitations and enhancements on applying the fair use doctrine for these situations.

By way of example, and particularly relevant to the discussion on orphan works and mass digitization, the Code of Best Practices for Academic and Research Libraries includes the following principle: “It is fair use create digital versions of a library’s special collections and archives and to make these versions electronically accessible in appropriate contexts.”⁴⁹ The Code of Best Practices suggests that several limitations to this broad principle exist, and:

[p]roviding access to published works that are available in unused copies on the commercial market at reasonable prices should be undertaken only with careful consideration, if at all. To the extent that the copy of such a work in particular collection is unique (e.g., contains marginalia or other unique markings or characteristics), access to unique aspects of the copy will be supportable under fair use. The presence of non-unique copies in a special collection can be indicated by descriptive entries without implicating copyright.⁵⁰

Other limitations referenced include taking “reasonable steps . . . to limit access to material likely to contain damaging or sensitive private attribution” and providing full attribution where reasonably possible.⁵¹ In addition to these limitations, the Code of Best Practices lays out several enhancements where libraries can bolster a fair use argument by, for example, “[a]dding criticism, commentary, rich metadata, and other additional value and context to the collection,”

Orphan Works for Libraries, Archives and Other Memory Institutions; Code of Best Practices in Fair Use for Poetry; Statement of Best Practices in Fair Use of Dance-Related Materials; Code of Best Practices in Fair Use for OpenCourseWare; Fair Use for Media Literacy Education; Code of Best Practices in Fair Use for the Visual Arts; and Best Practices in Fair Use for Documentary Filmmakers).

⁴⁹ *Id.* at 19-21.

⁵⁰ *Id.*

⁵¹ *Id.*

and where the collection is “appropriately publicized to scholars in the field and other persons likely to be especially interested.”⁵² It also notes that “[t]he fair use case 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,” among other enhancements.⁵³ These carefully articulated best practices provide useful guidance, particularly with the detailed limitations and enhancements, and are designed to provide the library community with greater comfort in relying on fair use.

Despite the comfort of the library community in relying on fair use to handle problems associated with orphan works or mass digitization, the Copyright Office report recommended legislative action and pilot programs on these issues. The Copyright Office’s proposal on mass digitization included an extended collective licensing proposal, to the surprise of many participants in the roundtables.⁵⁴ This proposal suggests an extended collective licensing pilot program would cover literary works; pictorial or graphic works published as illustrations, diagrams or similar adjuncts to literary works; and photographs.⁵⁵ It does not cover unique collections, such as those involving unpublished works, personal letters, or similar items.⁵⁶ As the Copyright Office explains in greater detail in its Notice of Inquiry soliciting comments on the proposed pilot:

⁵² *Id.*

⁵³ *Id.*

⁵⁴ HathiTrust, Mass Digitization Pilot Program; Request for Comments (Oct. 9, 2015), <http://copyright.gov/policy/massdigitization/comments/HathiTrust.pdf> [<https://perma.cc/9K4Q-RDX5>] (“At the roundtable I personally heard objections from creators, distributors, libraries and even potential collective management organizations. Our written comments noted that we ‘were struck by the degree to which participant stakeholders opposed the implementation of an extended collective licensing scheme in the United States, and believe that this demonstrates clearly that such a legislative solution should not be pursued.’ Thus, my colleagues and I were surprised to see the June Notice of Inquiry asking for still further comments on a voluntary pilot Extended Collective Licensing Program.”).

⁵⁵ U.S. COPYRIGHT OFF., *supra* note 41, at 104.

⁵⁶ *Id.*

Under the proposed framework, a collective management organization (CMO) representing copyright owners in a particular category of works would be permitted to seek authorization from the Register of Copyrights to issue licenses on behalf of both members and non-members of the CMO for certain mass digitization activities Once authorized, a CMO would be entitled to negotiate royalty rates and terms with users seeking to digitally reproduce and provide online access to a collection or body of copyrighted works for the benefit of the public, a community, or other specified users The CMO would be required to collect and distribute royalties to rightsholders within a prescribed period and to conduct diligent searches for non-members for whom it had collected payments. Copyright owners would have the right to limit the grant of licenses with respect to their works or to opt out of the system altogether.⁵⁷

The program would be limited to uses “undertaken for nonprofit or research purposes and without any purpose of direct or indirect commercial advantage.”⁵⁸ The Copyright Office’s proposed program also recommends inclusion of a fair use savings clause.⁵⁹

The creation of a collective management organization and registry, as well as its operation and management, comes with significant costs. The Copyright Office’s proposal fails to indicate who will bear such costs. While the Copyright Office seeks to address mass digitization through extended collective licensing, ultimately the proposed pilot program is an impractical solution with limited support. As the Library Copyright Alliance comments responding to the Copyright Office’s Notice of Inquiry notes, the proposed pilot program appears

⁵⁷ Mass Digitization Pilot Program; Request for Comments, 80 Fed. Reg. 110, 32614 (June 9, 2015), <http://www.copyright.gov/fedreg/2015/8ofr32614.pdf> [<https://perma.cc/PA8Q-FLEH>].

⁵⁸ U.S. COPYRIGHT OFF., *supra* note 41, at 89, 104; *see also* Mass Digitization Pilot Program; Request for Comments, *supra* note 57 (“Because the pilot is a limited project, such uses at this early juncture could be made only for nonprofit educational and research purposes and without any purpose of direct or indirect commercial advantage.”).

⁵⁹ U.S. COPYRIGHT OFF., *supra* note 41, at 105.

to be based off the Google Books Settlement,⁶⁰ overlooking the vast differences between Google Books and the Copyright Office's proposal.⁶¹ Under the Google Books Settlement, Google would have paid \$34.5 million to cover the start up costs of the registry system, would have paid out at least \$45 million for distribution to rightsholders, provided additional services beyond those that proposed in the pilot program, subsidized institutional subscriptions for libraries that partnered with Google in the project, and provided free public access terminals in public libraries.⁶² However, under the Copyright Office's proposal, "the institutional subscribers would bear the entire cost There would be no Google to subsidize libraries' purchase of institutional subscriptions or to pay the [collective management organization's] start-up costs."⁶³ The stark differences between the initial costs and subsidization of the Google Books Settlement versus the proposed pilot program will make it unlikely that the latter would succeed.

Aside from the enormous start up costs without any subsidies, collective management organizations are unlikely to actually ensure that rightholders are found and paid for the uses proposed in the pilot program. The costs of conducting searches for the rightholder could far exceed the license fees collected under the proposed program. The Authors Guild, which seeks to serve as a collective management organization under the Copyright Office's pilot program, even admits that searching for the correct rightholder comes at a heavy cost and could easily outweigh any license fees collected: "finding copyright owners from scratch can be quite costly. The required diligence, therefore, should bear some rational relation to the amount of money at stake. No one should spend \$100 to find an author owed \$5. There might be a threshold amount of money sent to authors to avoid

⁶⁰ *Id.* at 85. For more information regarding the Google Books Settlement model, see Jonathan Band, *The Book Rights Registry in the Google Books Settlement*, 34 COLUM. J.L. & ARTS 671 (2011).

⁶¹ Libr. Copyright All., Response to the U.S. Copyright Office's Notice of Inquiry on a Mass Digitization Pilot Program (Oct. 9, 2015), <http://www.librarycopyrightalliance.org/storage/documents/massdigitizationfinal.pdf> [<https://perma.cc/7NGQ-QLDT>].

⁶² *Id.* at 3.

⁶³ *Id.* at 4.

cutting checks for pennies.”⁶⁴ Furthermore, collective management organizations have a well-known history of corruption, mismanagement and lack of transparency.⁶⁵

Perhaps most importantly, it is difficult to see how the pilot program would succeed without the willingness of major institutions or organizations to purchase a license under the program.⁶⁶ As noted in Part IV.A, *infra*, institutions have already undertaken mass digitization projects for the purposes of preserving these important collections and also making them available to the public. Fair use already provides the solid legal footing for libraries to engage in digitization activities and it is unlikely that major research libraries would seek to take part in the burdensome and costly process proposed by the Copyright Office. In particular, the limitation of the proposed pilot program to non-profit, non-commercial educational and research purposes would require users such as libraries or educational institutions; this pilot program was not designed for corporate entities. As such, without the support of major universities or libraries, the program is unlikely to be successful.

IV. ARE FAIR USE SAVINGS CLAUSES ENOUGH?

Importantly, the Copyright Office proposes inclusion of a fair use savings clause.⁶⁷ While such recognition of fair use is welcome, the very existence of an extended collective licensing regime raises serious

⁶⁴ AUTHORS GUILD, COMMENTS OF THE AUTHORS GUILD: MASS DIGITIZATION PILOT PROGRAM, https://www.authorsguild.org/wp-content/uploads/2015/10/NOI_MassDigPilot_Authors-Guild.pdf [<https://perma.cc/5TDW-QH7Z>].

⁶⁵ See Jonathan Band & Brandon Butler, *Some Cautionary Tales About Collective Licensing*, 21 MICH. ST. INT’L L. REV. 687, 690 (2013).

⁶⁶ Libr. Copyright All., *supra* note 61; see also Triangle Res. Libr. Network, Comments of the Triangle Research Libraries Network in Response to the U.S. Copyright Office Notice of Inquiry Regarding a Mass Digitization Pilot Program (June 9, 2015), <http://www.copyright.gov/policy/massdigitization/comments/Triangle%20Research%20Libraries%20Network%20NRLN.pdf> [<https://perma.cc/6NPC-VPCF>] (“TRLN libraries are unlikely to purchase an ECL license offered by a CMO. These organizations would not provide an equitable, practical solution to rights issues connected with the material TRLN libraries deem most important to digitize . . . While we cannot speak for all libraries, we believe our digitization experiences and priorities are similar to those of many other U.S. research libraries.”).

⁶⁷ U.S. COPYRIGHT OFF., *supra* note 41, at 105.

concerns. First, users may choose to rely on licensing instead of exercising their fair use rights. The availability of a license may cause users to believe that they *must* purchase a license or they may choose to rely on a license to avoid the *potential* for litigation.

Even if a library believes the activity is fair use, counsel at the institution where the library is housed may advise purchasing a license to avoid the risk of litigation. Regardless of whether the activity would or would not be judged to be fair use by courts, the availability of a license could change the decision-making process at an institution. As comments by the University of California, Los Angeles Library note,

“We believe that mandating collective licensing is the wrong approach for both orphan works and mass digitization. Any system of compulsory licensing would undermine libraries’ legal entitlement to fair use. If a use is fair, then no permission or compensation should be required. There is a real and substantial risk that requiring payment of a license fee will discourage libraries from taking advantage of their fair use rights.”⁶⁸

In doing so, fair use jurisprudence may fail to move forward at its current speed. Furthermore, as the Authors Guild points out in comments to the Copyright Office, “If we had a collective licensing system in this country, there’s no doubt that courts would look at fair use differently in the mass digitization context.”⁶⁹ Indeed, courts have suggested that the existence of a license option would affect the fair use calculus.⁷⁰

⁶⁸ UCLA Libr., *supra* note 36.

⁶⁹ Authors Guild, *supra* note 64, at 4.

⁷⁰ American Geophysical Union v. Texaco, Inc., 60 F.3d 913, 930-31 (2d Cir. 1994) (“Despite Texaco’s claims to the contrary, it is not unsound to conclude that the right to seek payment for a particular use tends to become legally cognizable under the fourth fair use factor when the means for paying for such a use is made easier. This notion is not inherently troubling: it is sensible that a particular unauthorized use should be considered ‘more fair’ when there is no ready market or means to pay for the use, while such an authorized use should be considered ‘less fair’ when there is a ready market or means to pay for the use. The vice of circular reasoning arises only if the availability of payment is conclusive against fair use. Whatever the situation may have been previously, before the development of a market for institutional users to obtain licenses to photocopy articles, it is now appropriate to consider the loss of licensing revenues in evaluating ‘the effect of the

Not every use of copyrighted works requires the purchase of a license, even if one exists. Even if courts did not look at fair use differently in an environment of extended collective licensing, the fact that some might forego their fair use rights is detrimental to the development of fair use jurisprudence. As noted by Professor Peter Jaszi during a House Judiciary Subcommittee hearing on fair use, “Fair use, one might say, is like a muscle—it will grow in strength if it is exercised, and atrophy if it is not.”⁷¹ Without the robust exercise of fair use rights, the future of fair use may be jeopardized. The continuing development of fair use jurisprudence—particularly the more recent cases—strengthens the fair use right by helping to define appropriate fair uses, even where specific limitations and exceptions exist. In *HathiTrust*, the Second Circuit notably found HathiTrust’s activities regarding the creation and distribution of accessible format works to users with print disabilities to be fair use, making its determination solely on the grounds of fair use and without reference to whether the activities were permissible under Section 121 of the Copyright Act which specifically provides an exception to benefit those with disabilities.⁷² Thus, fair use is a right that exists in tandem with other rights to limitations and exceptions under the Copyright Act, but because of its flexible nature, it must be exercised to ensure that it is strengthened.

The Copyright Office’s proposal highlights just how detrimental the pilot program could be for fair use. It claims that the extended collective licensing regime would be limited to “where there is broad agreement that no colorable fair use claim exists” and “where the parties agree that a particular use would likely be deemed fair under established law, the portion of the license fee pertaining to that activity would likely be at or near zero.”⁷³ The problem with this

use upon the potential market for or value of journal articles. It is especially appropriate to do so with respect to copying of articles from *Catalysis*, a publication as to which a photocopying license is now available. We do not decide how the fair use balance would be resolved if a photocopying license for *Catalysis* articles were not currently available.”).

⁷¹ *The Scope of Fair Use: Hearing Before the Subcomm. on Courts, Intellectual Prop. and the Internet of the H. Comm. on the Judiciary*, 113th Cong. 10 (2014) (statement of Peter Jaszi, Professor, Faculty Director, Glushko-Samuelsan Intellectual Property Clinic, Wash. Coll. of L., American University), <https://judiciary.house.gov/wp-content/uploads/2016/02/113-82-86454-.pdf> [<https://perma.cc/5SZ8-BC6F>].

⁷² *Authors Guild v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014).

⁷³ U.S. COPYRIGHT OFF., *supra* note 41, at 101.

statement is that it suggests that parties must agree as to what is fair use, implying a user must first discuss and negotiate with a collective managing organization. Such an implication ignores the fact that fair use is a right and users engaging in this right need not engage in prior discussions or negotiations with rightholders. While there may be circumstances in which a user wishes to discuss a particular use with the rightholder, there is no notification requirement under the fair use statute.

Additionally, agreement as to what is fair use is unlikely and therefore, the Copyright Office's suggestion that extended collective licensing would take place in more limited circumstances is incorrect. The recent litigation between the Authors Guild and HathiTrust highlights this disagreement. Despite a clear fair use savings clause under Section 108, which governs specific exceptions for libraries and archives,⁷⁴ the Authors Guild still tried to advance an argument that HathiTrust's activities went beyond the scope of permissible activity under Section 108 and therefore could not be fair use. While the Second Circuit dismissed this argument, relegating discussion of this issue to a mere footnote, other rightholders may try to make similar arguments to restrict the use of fair use.

Even in situations where precedent strongly favors an understanding that a particular use would be considered a fair use—such as in *HathiTrust*—rightholders might try to argue that the precedent is wrongly decided and that recent jurisprudence has applied fair use too broadly.⁷⁵ They might also attempt to distinguish precedent on the basis of minor factual differences. As a result, agreement as to what is fair use may not be as simple as the Copyright Office suggests in its proposal.

⁷⁴ 17 U.S.C. § 108(f)(4) (2005).

⁷⁵ See sources cited *supra* note 40. In fact, outlandish statements have already been made with respect to recent fair use jurisprudence. In the orphan works/mass digitization roundtables hosted by the Copyright Office in March 2014, one participant compared fair use case law to *Plessy v. Ferguson*, the 1892 Supreme Court case that upheld the “separate but equal” doctrine until it was overturned by *Brown v. Board of Education* in 1954.

V. LIBRARY ACTIVITY TO PROMOTE ACCESS TODAY

A. Digitization of Special Collections

Despite supposedly limiting extended collective licensing pilot program to areas where “no colorable fair use claim exists,” the Copyright Office uses the example of “depression era photographs” as a type of project that would benefit from the pilot program.⁷⁶ This suggestion highlights the discrepancy between what is considered fair use by libraries—and supported by fair use jurisprudence—and what the Copyright Office believes to be fair use. A collection of depression era photographs is likely to represent works that are out-of-commerce and largely orphaned. It is unlikely that they rightholders for most depression era photographs could be identified and located; they are unlikely to be exploited commercially and digitizing and making them accessible online would therefore not harm the original market for the works. They are likely to be a part of a special collection that libraries would feel comfortable digitizing and providing access for the public, and would likely include enhancements that make the collection more useful, such as the inclusion of metadata.⁷⁷

The New York Public Library’s (NYPL) digitization of a collection of materials from the 1939 New York World’s Fair provides a perfect counterpoint to the Copyright Office’s suggestion that an extended collective licensing regime is necessary. After receiving 2,500 boxes of records and documents, as well as 12,000 promotional photographs,

⁷⁶ *Thoughts on Fair Use and the Copyright Office Report/Proposal on Mass Digitization*, ARL POLICY NOTES, <http://policynotes.arl.org/?p=1102> [<https://perma.cc/BGC7-T65F>] (“the Office is particularly interested in stakeholder views regarding examples of mass digitization projects that may be appropriate for licensing under the proposed pilot. These comments may include (but need not be limited to) descriptions of particular collections of copyrighted works (e.g., Depression-era photographs) that prospective users may wish to digitize and make available through ECL”).

⁷⁷ The ARL Code of Best Practices provides that, as a principle, it is fair use to digitize and make electronically available special collections in appropriate contexts. It also provides a number of limitations, but also enhancements. For example, “The fair use case 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.” Another “enhancement” provides that “Adding criticism, commentary, rich metadata, and other additional value and context to the collection will strengthen the fair use case.” ASS’N OF RES. LIBR., *supra* note 47.

NYPL moved forward in digitizing and making available⁷⁸ these works after conducting a fair use analysis.⁷⁹ If NYPL had conducted the project under an extended collective licensing regime, the project would have been more burdensome and would limit the access of these otherwise unused works. Although the Copyright Office's proposal has the fair use savings clause, the fair use decision-making process could be different in some institutions if an extended licensing program were available. As noted in comments by UCLA Libraries:

Any system of compulsory licensing would undermine libraries' legal entitlement to fair use. If a use is fair, then no permission or compensation should be required. There is a real and substantial risk that requiring payment of a license fee will discourage libraries from taking advantage of their fair use rights. Further, since by definition copyright holders cannot be identified or located for orphan works, there is no rightful person or entity able to accept the collected fees.⁸⁰

The existence of an unnecessary extended collective licensing system could hamper the exercise of fair use, even with the inclusion of a savings clause.

NYPL is not unique in its reliance on fair use. In comments to the Copyright Office, numerous libraries and consortia described their mass digitization projects. Boston College digitized a collection of records of former Speaker of the House, "Tip" O'Neill.⁸¹ The University of Massachusetts, Amherst digitized a collection of W.E.B. DuBois' letters, essays, lectures, fiction and non-fiction writing, research notes and photographs.⁸² George Mason University holds the

⁷⁸ N.Y. Pub. Libr., *supra* note 36.

⁷⁹ *Id.*

⁸⁰ UCLA Libr., *supra* note 36.

⁸¹ Boston Libr. Consortium, Mass Digitization Pilot Program: Request for Comment (Oct. 8, 2015), <http://blogs.umass.edu/lquilter/files/2015/10/BostonLibraryConsortium-20151009-NOIMassDigitization-FINAL-signed.pdf> [<https://perma.cc/U8Y6-2ZF8>].

⁸² *Id.* at 4.

Arthur Scott Photograph Collection,⁸³ which includes photographs taken by Arthur Scott between 1939-1974, as well as other photographs and correspondence.⁸⁴ Emory put together the EU Pix Collection,⁸⁵ including a large mix of items related to Emory and its history.⁸⁶ The University of North Carolina at Chapel Hill digitized the Frank Porter Graham papers collection⁸⁷ and a collection of historic North Carolina postcards.⁸⁸ Under the pilot program, these digitization projects would be “completely infeasible.”⁸⁹

Extended collective licensing is an inappropriate solution to special collections. Not only is digitization of these types of works a fair use, going forward with collective licensing could come at a great cost to the institution given the sheer volume of works in the collection. Libraries have tended to focus digitization efforts on special collections; in other words, “on material that is not currently discoverable and on material that is unique, rare, or ephemeral.”⁹⁰ Additionally, given the likely orphan works status of these types of works, it is unlikely that a collective management organization would be able to locate the actual rightholder. Thus, any funds collected under a pilot program would unlikely be paid to any actual

⁸³ *Guide to the Arthur E. Scott photograph collection, 1910-1976*, GEORGE MASON U., http://sca.gmu.edu/finding_aids/scotta.html [<https://perma.cc/82ZU-ELAC>].

⁸⁴ ASS'N OF S.E. RES. LIBR., COMMENTS OF THE ASSOCIATION OF SOUTHEASTERN RESEARCH LIBRARIES IN RESPONSE TO THE UNITED STATES COPYRIGHT OFFICE'S ORPHAN WORKS AND MASS DIGITIZATION: NOTICE OF INQUIRY (Sept. 1, 2015), http://www.aserl.org/wp-content/uploads/2015/09/ASERL_Response_USCO_Mass-Dig_Orphan_Works_Rpt.pdf [<https://perma.cc/YX5U-YADY>].

⁸⁵ *Emory University Photograph Collection 1860-2003*, EMORY U., <https://findingaids.library.emory.edu/documents/eua0111eupixx/> [<https://perma.cc/UPR8-5UCD>].

⁸⁶ *Id.*

⁸⁷ *Frank Porter Graham Papers, 1908-1990*, U. N.C., http://www2.lib.unc.edu/mss/inv/g/Graham,Frank_Porter.html [<https://perma.cc/HN6M-LQRH>].

⁸⁸ *North Carolina Postcards*, U. N.C., http://www2.lib.unc.edu/dc/nc_post/index.php [<https://perma.cc/T5L6-4MC6>].

⁸⁹ Boston Libr. Consortium, *supra* note 81, at 4.

⁹⁰ Triangle Res. Libr. Network, *supra* note 66.

rightholders and are more likely to be held by the collective management organization.

As a result of an extended collective licensing regime, not only would funds be unlikely to reach the rightholder, access to large collections would be restricted because of an inability of libraries to digitize them due to the heavy burdens and high costs. Researchers would be forced to travel to the library holding the collection to conduct research, thus “curtail[ing] American scientific and commercial innovation, limit[ing] our cultural and historical insights, and inhibit[ing] the growth of artistic knowledge. The results are incalculable.”⁹¹ An ECL regime “would have a profound chilling effect on the efforts by libraries and other cultural heritage organizations to improve public access to their collections for research, teaching, and learning.”⁹²

Fair use provides a far better path for the digitization of special collections than extended collective licensing regimes. Indeed, as comments by the Boston Library Consortium point out:

It is no accident that the case law regarding copyright infringements by libraries and archives is exceedingly sparse, and virtually non-existent with respect to special collections. It is in fact a testament to the careful attention to the equities that librarians and archivists provide, as well as a reflection of the obvious fact that for many ‘orphan works,’ no claimants exist.⁹³

Similarly, the Triangle Research Libraries Network point out that, “We have digitized countless in-copyright works, either with

⁹¹ ASS’N OF S.E. RES. LIBR., *supra* note 84.

⁹² *Id.*

⁹³ Boston Libr. Consortium, *supra* note 81; *see also* ASS’N OF S.E. RES. LIBR., *supra* note 84 (“ASERL members remain focused on the underlying purpose of copyright, ‘to promote the progress of Science and useful Arts, by securing for a limited Time to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.’ We believe libraries should continue to rely on the well-established principles of fair use as announced in the Copyright Act, 17 U.S.C. §107. These principles have served libraries and interested copyright holders remarkably well over time. The current legal framework provides a fair and equitable means for resolving disputes relating to works that have been digitized and made available on the Internet. Fair use has survived multiple legal challenges, ensuring libraries and other cultural heritage organizations can continue to provide improved access to the nation’s intellectual and cultural output.”).

permission from rightsholders or under the doctrine of fair use or other legal defenses, coupled with a takedown mechanism, should that prove necessary. From among our vast holdings and digitization projects, we are unable to identify a single instance in which we would benefit by pursuing a license under an extended collective licensing (ECL) system”⁹⁴

B. *Beyond Print*

Beyond the mass digitization of texts, such as the corpus of *HathiTrust*, or printed materials such as photographs, documents or brochures, fair use jurisprudence supports the digitization and copying of other materials⁹⁵ and libraries are already engaged in such activities. The University of North Carolina, for example, has a unique collection of North Carolina films chronicling aspects of life in the state during the twentieth century.⁹⁶ It also digitized rare sound recordings and motion pictures for its Southern Folklife Collection.⁹⁷ Duke University digitized historic TV commercials in its collection, *ad Views*, after securing agreements from many rightsholders and also relying on fair use.⁹⁸

C. *Accessibility*

In addition to digitizing collections for use by general users, libraries are also heavily invested in promoting accessibility of their collections to those with disabilities. The ability to digitize materials has revolutionized the ability to provide access to works for persons who are blind, visually impaired or otherwise print disabled. Mass

⁹⁴ Triangle Res. Libr. Network, *supra* note 66.

⁹⁵ See *Fox v. TVEyes*, 43 F.Supp. 3d 379 (S.D.N.Y. 2014).

⁹⁶ Triangle Res. Libr. Network, *supra* note 66.

⁹⁷ *Id.*; see also *Southern Folklife Collections Receives \$986,000 Grant from Andrew W. Mellon Foundation*, UNC LIBR., <http://blogs.lib.unc.edu/news/index.php/2015/06/sfc-mellon-grant/06/sfc-mellon-grant/> [<https://perma.cc/SD68-S3X2>].

⁹⁸ *Thoughts on Fair Use and the Copyright Office Report/Proposal on Mass Digitization*, *supra* note 76 (Duke enhanced its fair use position by adding additional videos to the collection featuring executives talking about TV advertising in the early 1960s as well as faculty members discussing the ways they used the materials in teaching).

digitization has provided immense benefits by allowing persons with print disabilities to have accessible formats of these digitized works. As noted by a representative of the National Federation of the Blind, extended collective licensing regimes could have a chilling effect on mass digitization efforts and, as a result, limit access to works for persons with visual impairments.⁹⁹ Libraries have long been committed to promoting accessibility and have created and distributed accessible formats to patrons with print-disabilities under both the Chafee Amendment,¹⁰⁰ the specific exception governing creation of accessible format works, and fair use.¹⁰¹

For example, HathiTrust Digital Library, a partnership of more than 100 academic and research institutions, has provided accessible formats of digitized books to users who are blind or print-disabled¹⁰²—allowing students with print disabilities at member schools to have access to these works¹⁰³—and recently announced a new collaboration with the National Federation of the Blind.¹⁰⁴ This partnership opened up more than 14 million digital books in the HathiTrust repository to persons with print disabilities.¹⁰⁵

One issue that has been a priority for libraries for many years is advancement of the first World Intellectual Property Organization

⁹⁹ NAT'L FED'N OF THE BLIND, ORPHAN WORKS AND MASS DIGITIZATION: REPLY COMMENTS TO NOTICE OF INQUIRY AND RESPONSES (Mar. 6, 2013), http://www.copyright.gov/orphan/comments/noi_11302012/National-Federation-of-the-Blind.pdf [<https://perma.cc/J53C-RGXN>].

¹⁰⁰ 17 U.S.C. § 121 (2004).

¹⁰¹ 17 U.S.C. § 107; *see also* Authors Guild v. HathiTrust, 755 F.3d 87 (2d. Cir. 2014).

¹⁰² *HathiTrust at U-M, NFB to make 14M+ books accessible to blind and print-disabled users*, HATHITRUST (June 29, 2016), https://www.hathitrust.org/hathitrust_NFB_announcement [<https://perma.cc/EC7K-S6CF>] [hereinafter *HathiTrust*] (“Supporting print-disabled users has been a focus of HathiTrust since the very beginning, and we have long provided students at HathiTrust member schools with access to our collection’ said Mike Furlough, executive director of HathiTrust. ‘The collaboration with NFB is an important turning point, because we are now striving to help non-academic print-disabled users for the first time.’”).

¹⁰³ *Accessibility*, HATHITRUST, <https://www.hathitrust.org/accessibility> [<https://perma.cc/Z9L6-PU6F>].

¹⁰⁴ *HathiTrust*, *supra* note 102.

¹⁰⁵ *Id.*

(WIPO) treaty designed for the benefit of the users rather than creators, specifically one for the benefit of persons who are print disabled. In June 2013, WIPO concluded the Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled (Marrakesh Treaty), a treaty with minimum standards for limitations and exceptions for the creation and distribution of accessible formats and cross-border sharing of these formats.¹⁰⁶ The Marrakesh Treaty was designed to alleviate a problem known as the “book famine” where “only a small fraction of published books—estimated at less than 7 percent—are made in accessible formats. This percentage is even smaller in low-income countries, resulting in a ‘book famine’ where persons who are print disabled have no access to the vast majority of works.”¹⁰⁷ Allowing cross-border sharing should help grow collections of accessible format works for countries that share a common language and avoid unnecessary duplication of efforts in the creation of these formats. The Marrakesh Treaty was adopted on June 27, 2013 and will go into force on September 30, 2016, following ninety days after the twentieth ratification of the Marrakesh Treaty. The first twenty ratifications of the Marrakesh Treaty, in order of ratification or accession, include: India, El Salvador, United Arab Emirates, Mali, Uruguay, Paraguay, Singapore, Argentina, Mexico, Mongolia, Republic of Korea, Australia, Brazil, Peru, Democratic People’s Republic of Korea, Israel, Chile, Ecuador, Guatemala and Canada.¹⁰⁸

Although the overall percentage of accessible format works is low in the United States, estimated at no more than five percent,¹⁰⁹ the

¹⁰⁶ *Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled*, WIPO, <http://www.wipo.int/treaties/en/ip/marrakesh/> [<https://perma.cc/9VQC-7XGK>].

¹⁰⁷ JONATHAN BAND & KRISTA COX, U.N. DEV. PROGRAMME, OUR RIGHT TO KNOWLEDGE: LEGAL REVIEWS FOR THE RATIFICATION OF THE MARRAKESH TREATY FOR PERSONS WITH PRINT DISABILITIES IN ASIA AND THE PACIFIC (2015), http://www.asia-pacific.undp.org/content/dam/rbap/docs/Research%20&%20Publications/hiv_aids/rbap-hhd-2015-our-right-to-knowledge.pdf [<https://perma.cc/7WM6-2N6W>].

¹⁰⁸ *WIPO Administered Treaties: Notifications: Marrakesh VIP Treaty*, WIPO, http://www.wipo.int/treaties/en/ShowResults.jsp?search_what=N&treaty_id=843 [<https://perma.cc/HNE2-X6ZH>].

¹⁰⁹ *Marrakesh Treaty to Facilitate Access to Published Works for Person Who Are Blind, Visually Impaired or Otherwise Print Disabled*, NAT’L FED’N OF THE BLIND, <https://nfb.org/books-without-borders> [<https://perma.cc/UKW7-XJWP>].

United States does have one of the largest collections of accessible formats worldwide. Libraries in the United States have therefore been working with key stakeholders, including groups representing the blind and visually impaired community, to advance ratification efforts domestically. Ratification of the Marrakesh Treaty will provide greater certainty for libraries to fulfill their missions of providing access, particularly in sharing these accessible formats with beneficiaries in other countries.

Even beyond access for those with print disabilities, libraries view achieving better accessibility across the spectrum of disabilities as a priority. For example, libraries work to ensure that those with hearing impairments have the accessible formats that they need, such as appropriately captioned video materials. As technology continues to evolve and digital resources are more easily shared than hard copy resources, libraries can and should work together to avoid wasted duplication and efforts in promoting full accessibility to their patrons.

D. Preservation of Websites and 21st Century Ephemera

While libraries have always committed to preservation activities to ensure the protection of the historical and cultural record for centuries to come, the digital age provides new challenges and opportunities. Preservation by libraries and other cultural heritage institutions is essential because without such activities, much of the historical record would be lost as publishers and other rightholders may not have the interest or incentive to preserve all of their works.¹¹⁰

The advances of technology, particularly over the last twenty years, have supported the vast dissemination of information and culture. Technology has truly changed communication systems, allowing easy dissemination of information through websites and blog posts and user created videos and sound recordings. These websites, videos and other materials truly represent a part of today's history and culture but "[o]ne 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

¹¹⁰ See, e.g., LIBR. COPYRIGHT ALL., COMMENTS ON MANDATORY DEPOSIT OF ELECTRONIC BOOKS AND SOUND RECORDINGS AVAILABLE ONLY ONLINE (Aug. 18, 2016), http://www.librarycopyrightalliance.org/storage/documents/MandatoryDepositNOI_final.pdf [<https://perma.cc/DB9P-K249>].

media become obsolete.”¹¹¹ A 2013 study highlighted the problem of “link-rot” finding that approximately fifty percent of links in United States Supreme Court opinions did not link to the originally cited materials; more than percent of links used in citations in the Harvard Law Review and other journals failed to link to the original material.¹¹²

While the ease of creating a website to disseminate information is a welcome advance in technology, without preservation efforts by libraries and other memory institutions, these records may not last five years into the future, much less five hundred years from now. As a result of the increasingly ephemeral nature of history, culture and scholarly resources in the digital age, libraries have moved to preserve these works. At Columbia University, for example, digital preservation activities have included several significant collections:

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 minidisks, 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. Minidisks 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

¹¹¹ PRESERVATION AND REUSE OF COPYRIGHTED WORKS: SUBCOMM. ON COURTS, INTELLECTUAL PROP. AND THE INTERNET OF THE H. COMM. ON THE JUDICIARY, 113TH CONG. 3 (2014) [hereinafter *Neal Hearings*] (statement of James G. Neal), <http://www.librarycopyrightalliance.org/storage/documents/testimony-jim-neal-2apr2014.pdf> [<https://perma.cc/JSY9-5EMB>].

¹¹² Jonathan L. Zittrain, et. al., *Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations*, 127 HARV. L. REV. F. 176 (2014).

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 2008, 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.¹¹³

Without these digital preservation efforts, these important records—such as those related to 9/11, pre-World War II experiences and customs, and human rights concerns—could be lost forever.

¹¹³ *Neal Hearings*, *supra* note 111, at 4.

While digital resources may not deteriorate in the same manner as a paper-based book or manuscript, these materials must still be preserved because of their ephemeral nature and the risk that particular media or technologies will become obsolete. Libraries can and should ensure that the cultural record is protected and can do so through fair use. As noted by James G. Neal before the House Judiciary Subcommittee:

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.¹¹⁴

Although the current Copyright Act was largely created in the pre-digital era, because of the flexible nature of the fair use doctrine, these preservation activities are still possible. Fair use sufficiently updates the Copyright Act because its flexible nature ensures that it can adapt to new technologies and does not require legislative amendment each time a new technology or use is discovered. Thus, while preservation today may differ from the analog world, libraries are still able to continue their missions and preserve material in the digital world because of fair use.

E. Institutional Repositories

Another mechanism for promoting preservation, discovery and learning takes place through institutional repositories run by research libraries. These repositories—many of which were designed to promote open access—can ensure that articles and manuscripts deposited are preserved and provide access to these materials.¹¹⁵

¹¹⁴ *Id.*

¹¹⁵ See, e.g., Ellen Finnie & Greg Eow, *Beware the Trojan Horse: Elsevier's repository pilot and our vision for IRs and Open Access*, IN THE OPEN (May 31, 2016), <http://intheopen.net/2016/05/beware-the-trojan-horse-elseviers-repository-pilot-and-our-vision-for-irs-open-access/> [<https://perma.cc/BK85-W772>] (Noting that the larger

Libraries have strongly supported open access because open access models promote access to knowledge. Open access supports opportunities for readership by lowering costs for the readers and can increase visibility of materials both inside and outside of the academy. Scholars often support open access because of the great contributions to discussions, and some studies show that citations to open access materials take place both earlier and more frequently.¹¹⁶ Numerous institutions have open access policies,¹¹⁷ ranging from broad policies in which faculty members grant their university a non-exclusive license to exercise the rights to the copyright to the faculty's scholarly articles¹¹⁸ to simply encouraging faculty to use open access or an opt-in approach.¹¹⁹ At many institutions, the open access policy is driven by the library and, at some, the extent of the policy applies only to the libraries or specific departments.¹²⁰

Research funders have recognized the value of sharing information and knowledge and, therefore, the value of public access or open access resources. As a result, funders have begun to incorporate public or open access requirements in their grants. The United States federal government, for example, has set forth public access requirements because it "funds research with the expectation that new ideas and discoveries from the research will propel science,

objectives include "digital preservation, long-term access, and access to the fruits of scholarship as democratizing knowledge and promoting social justice.").

¹¹⁶ ALMA SWAN, *THE OPEN ACCESS CITATION ADVANTAGE* 17 (2010), http://eprints.soton.ac.uk/268516/2/Citation_advantage_paper.pdf [<https://perma.cc/3D8Z-TZJX>].

¹¹⁷ *Open Access Policies at Other Universities*, MIT LIBR., <https://libraries.mit.edu/scholarly/mit-open-access/open-access-at-mit/mit-open-access-policy/mit-faculty-open-access-policy-faq/other-university-policies/> [<https://perma.cc/29ZN-X9CP>].

¹¹⁸ Examples include Harvard University, the Massachusetts Institute of Technology, and University of California System.

¹¹⁹ *See, e.g.*, S. ILL. U. EDWARDSVILLE, *OPEN ACCESS POLICY*, <http://www.siue.edu/lovejoylibrary/pdf/OAPolicy.pdf> [<https://perma.cc/A7CJ-P7YU>]; Faculty Open-Access Statement of Principles, U. PENN., <http://www.upenn.edu/almanac/volumes/v58/no3/openaccess.html> [<https://perma.cc/5KPN-LYCB>].

¹²⁰ *Open Access Policies at Other Universities*, *supra* note 117 (Examples of universities where the policy applies only to the library include Arizona State University, Miami University of Ohio, and Wake Forest University).

stimulate the economy, and improve the lives and welfare of Americans.”¹²¹ Following a February 22, 2013 memorandum from the White House Office of Science and Technology Policy directing federal funding agencies with research budgets of more than \$100 million to develop plans to support increased public access,¹²² federal agencies have issued guidance and plans on establishing public access.¹²³ Foundations, such as the Wellcome Trust¹²⁴ and Bill and Melinda Gates Foundation¹²⁵ have similarly implemented open access policies. Libraries play a critical role in supporting faculty in understanding and ensuring compliance with these policies, as well as promoting open access (and thus preservation of these materials) through their repositories.

As more research libraries build institutional repositories designed to facilitate open access and greater collaboration, knowledge will not only be preserved for future generations, but today’s scholars can take advantage of these systems to better conduct research. Institutional repositories and shared databases allow research, data and scholarly activities to become more discoverable and promote the detection of links that would not be possible in the individual context.

¹²¹ *Focus Areas: Public Access Policies*, ASS’N OF RES. LIBR., <http://www.arl.org/focus-areas/public-access-policies#.V73IGz4rJNo> [<https://perma.cc/W8FU-55XL>].

¹²² John P. Holdren, *Expanding Public Access to the Results of Federally Funded Research*, WHITE HOUSE (Feb. 22, 2013), <https://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research> [<https://perma.cc/7H5S-N3WR>].

¹²³ *White House Directive on Public Access to Federally Funded Research Data*, ASS’N OF RES. LIBR., <http://www.arl.org/focus-areas/public-access-policies/federally-funded-research/2696-white-house-directive-on-public-access-to-federally-funded-research-and-data#.V73Igj4rJNo> [<https://perma.cc/TP6V-TLE5>].

¹²⁴ *Open access policy*, WELLCOME, <https://wellcome.ac.uk/funding/managing-grant/open-access-policy> [<https://perma.cc/4FZ4-KQ3P>].

¹²⁵ *Open Access Policy*, BILL AND MELINDA GATES FOUND., <http://www.gatesfoundation.org/How-We-Work/General-Information/Open-Access-Policy> [<https://perma.cc/NFX4-3NY4>] (requiring that publications be discoverable and accessible online, including deposit in specified repositories with proper metadata tagging; requiring publication under the Creative Commons Attribution 4.0 generic license (CC-BY 4.0); requiring immediate open access without any embargo period; requiring all data underlying the published research to be made open immediately).

VI. CONCLUSION

Libraries today in many ways play the same role they always have: finding ways to preserve cultural and historical heritage and promoting access to information and knowledge. The evolution of technology has presented both new challenges and new opportunities for libraries to achieve their missions. These new technologies have led to a more rapid pace of dissemination of information and creation of new forms of communications which make libraries essential in the digital age. While some portions of the copyright law may not have been written for the digital environment, the fair use doctrine provides a sufficient basis for libraries to continue to preserve, share and disseminate knowledge and information.