



2017

Piece by Piece Review of Digitize-and-Lend Projects Through the Lens of Copyright and Fair Use

Michelle M. Wu

Georgetown University Law Center, mmw84@law.georgetown.edu

This paper can be downloaded free of charge from:
<https://scholarship.law.georgetown.edu/facpub/1974>
<https://ssrn.com/abstract=2968410>

Legal Reference Services Q. (forthcoming 2017)

This open-access article is brought to you by the Georgetown Law Library. Posted with permission of the author.
Follow this and additional works at: <https://scholarship.law.georgetown.edu/facpub>

Piece by Piece Review of Digitize and Lend Projects Through the Lens of Copyright and Fair Use

Michelle M. Wu¹

Digitize-and-lend library projects² can benefit societies in multiple ways, from providing information to people in remote areas, to reducing duplication of effort in digitization, to providing access to people with disabilities. Such projects contemplate not just digitizing library titles for regular patron use, but also allowing the digitized versions to be used for interlibrary loan (ILL), sharing within consortia, and replacing print copies at other libraries. Many of these functions are already supported within the analog world (e.g., ILL), and the digitize-and-lend concept is largely a logical outgrowth of technology, much like the transitioning from manual hand duplication of books to printing presses. The purpose of each function is to facilitate user access to information. Technology can amplify that access, but in doing so, libraries must also be careful not to upset the long established balance in copyright, where authors' rights sit on the other side of the scale from public benefit.³

This article seeks to provide a primer on the various components in a digitize-and-lend project, explore the core copyright issues in each, and explain how these projects maintain the balance of copyright even as libraries take advantage of newer technologies.

Benefits of Digitize and Lend

What is digitize-and-lend? There are several activities that fit that description, including these three models:

- Digitizing a print title for preservation purposes and lending only physical facsimiles created from the resulting digital copy.
- Digitizing a print title and lending or giving the e-copy to visually disabled patrons only.
- Digitizing a print title to create an e-equivalent, where the e-book is intended for regular patron use.

¹ Associate Dean for Library Services and Professor of Law, Georgetown University Law Center

² For a world-wide model, see Open Library at <https://openlibrary.org/about>. For descriptions on how a consortia or shared model might work within the United States, see Michelle M. Wu, *Building a Collaborative Digital Collection, a Necessary Evolution in Libraries*, 103 L. LIBR. J. 527 (2011).

³ U.S. CONST. Art. I, Sec. 8, Cl. 8 empowers Congress "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." Both leading copyright experts and courts have routinely read this to mean that "...encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors." (Mazer v. Stein, 347 U.S. 201, 219 (1954)). See also MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT, §1.03.

For the purposes of this paper, “digitize-and-lend” refers only to the third model above, and to only one variation of that model. That variation assumes that any title digitized is still protected by copyright,⁴ and that a library will only digitize a title that has been legitimately acquired (e.g., purchase or gift). A library circulating a digitized title may only use simultaneously the number of copies that it has legitimately acquired. For example, if it owns three copies of a title and digitizes one copy, it could choose to circulate one digital copy and two print, or three digital copies, or two digital copies and one print; in all cases, it could only circulate the same number of copies that it owned before digitization. Circulation in any format is controlled, whether through an integrated library system or an e-reader (e.g. Adobe Digital Editions) that recognizes digital rights so that only one user can use any given copy at a given time. Last, libraries using this model will serve only patrons within the boundaries of the United States and its territories.

The other two models are not discussed, as they are more straightforward and do not raise the types of complex copyright questions that surround the third. And variations of the third are not explored, as they are uncommon and are heavily dependent on facts of their specific implementations.

Before examining the component pieces of a digitize-and-lend project, it is useful to understand the benefits of this approach and why libraries might adopt it as a core mechanism to accomplish their service missions. Those benefits generally fall into three different categories: cost savings, access to a broader range of materials by all participating libraries, and service to underserved populations.

Cost savings come in multiple forms. The first is in staff time for interlibrary loans, collection development, and digitization. As digital items can be loaned without packaging and shipping, staff time is saved on those activities. Similarly, if the project involves multiple libraries, collection development is shared --- even if only for the parts of the collection that are not regularly used --- the review and selection of such materials need not be duplicated at every library; subjects could be assigned to member libraries for a more complete review and the building of a more comprehensive collection. Staff also need not spend time on recalls, holds, and searching for claimed-returned items, as returns and recalls could be automated; recalcitrant patrons could not retain a title for longer than allowed.

Last, if libraries with common titles can digitize the title once and provide it to others for their collections (where the receiving libraries hold the print title), then the libraries holding that same title need not all digitize the title independently. Libraries collectively can digitize one copy of each title and use it to replace all of their print equivalents, if desired. The library undertaking the digitization would see increased costs, but the receiving libraries would have

⁴ Works in the public domain may be freely reproduced, shared, performed, and displayed, and therefore fall outside of the scope of this article.

none. While this is not necessarily a cost savings in the same manner as the others described above, it does ensure that labor is not duplicated by digitizing the same title repeatedly.

There could also be savings in collection costs, as libraries will no longer need to purchase the same content in different formats as technologies fall out of use or style. If a library is interested in the *Wizard of Oz*, they need not replace a film version with a VCR version, then a DVD version, then a streaming version, unless they want added content (e.g., colorization, outtakes) that may exist on newer releases. Those libraries satisfied with the content of the original can simply shift that original format into a current or more usable format and discontinue use of the older format. For public libraries, this is particularly compelling, as the best and most effective use of public funds would be to purchase as much unique content as needed for the community, not to spend funds repeatedly on identical content where multiple copies are not needed.

Other less meaningful cost savings can also be achieved through reduction of operating costs. If libraries no longer need to ship books, they save both on shipping and packaging, and the monies for these operations can be redirected into collection purchases. They also need not worry about security measures such as tattle tape or RFID tags for digital items, thereby reducing the costs not only of the materials for each book, but also potentially reducing the need for security gates or security guards for physical collections. Security services may still be necessary for library spaces or other uses, but they would reduce the costs attendant to the protection of print materials. On the flip side, there may be reduced fine revenues from overdue books, but that could be offset by the cost savings of staff time, not to mention the non-financial aspect of having a title available to users in a timely manner.

The other major benefit is in the expansion of access, and this also comes in various forms: access to a broader range of materials, more user-friendly materials, and service to underserved populations.

As budgets have declined and publishing has proliferated,⁵ many academic libraries have reduced the active collection of research titles in areas less frequently used by their user populations, choosing to acquire these only when requested. Logically, this means that the collections of the libraries of the 21st century are overall less complete in published content than the libraries of earlier years.⁶ In terms of immediate availability of recently published information, this may not have a significant impact, so long as they remain in print. But once the title is no longer carried by the publisher, options for discovery shrink considerably. By coordinating selection and purchase of these titles by libraries, libraries ensure that a wider and more comprehensive range of titles on any topic remain accessible to future generations. This

⁵ Steve Coffman, *How Low Can Our Book Budgets Go?* at <https://americanlibrariesmagazine.org/2013/10/14/how-low-can-our-book-budgets-go/>

⁶ Michael Levine-Clark, Margaret Jobe & Sara Holladay, *Uniqueness and Collection Overlap in Academic Libraries*, Proceedings of the Charleston Conference (2009) at <http://docs.lib.purdue.edu/charleston/2009/ContentDev/12/>

will enable scholars at all participating institutions to access comprehensive collections instead of only the titles that they are able to identify and request themselves, leading to more informed discourse. Libraries also increase the chances of survival of a “last copy” should disaster strike the few, or single library, owning a title, something that cannot be done effectively in print.

The second benefit in access is simply usability. Libraries with large microform collections or with collections of non-print materials (e.g., A/V) regularly note user dissatisfaction with these formats. Users simply avoid microform if at all possible,⁷ and A/V materials can be difficult to access if the right equipment cannot be found or is not operational at the time the resource is needed. By digitizing a title and continually converting it to a format that is easily accessed through library machinery or equipment commonly used by the population, the title will be easily accessed throughout its lifespan.

The last benefit speaks to populations for which distance is a barrier to access. The Pew Research Center conducted a study on public library use and its data showed that individuals living in rural areas or who had physical disabilities reported greater difficulties in accessing their libraries’ resources than other populations.⁸ In converting materials from a physical form to a digital one, these barriers can be overcome for many patrons. A book can be checked out even when a library is closed or when a patron is unable to physically visit their library. A user does not have to lose time commuting to a library to access its materials. Users themselves may also see savings or value in accessing information that might otherwise be available only in remote locations or deemed so fragile that they cannot be handled in their original formats.

With all of these benefits, it is not surprising that libraries would view digitize-and-lend projects as a tool to enhance the services and collections that they provide. But regardless of how beneficial an action might be, it also must pass the legal tests set by copyright laws.

Copyright Infringement and the Affirmative Defenses of First Sale and Fair Use

From its first appearance in United States’ federal law, copyright has sought to balance the interests between authors and the public.⁹ The floor debates and congressional documents accompanying every major piece of copyright legislation show this constant struggle as well as the unwavering commitment to both the public interest and authors’ rights.

Given those interests, it should be unsurprising that copyright laws are fairly evenly divided across these interests. Section 106 reserves to the copyright owner broad control over

⁷ See Stephen R. Salmon, *User Resistance to Microforms in the Research Library*, 3 MICROFORM & DIGITIZATION REV. 194 (1974) and Colleen Cook & Fred M. Heath, *User’s Perception of Library Service Quality: A LibQUAL+ Qualitative Study*, 49 LIB. TRENDS 548 at 573 (2001).

⁸ Kathryn Zickuhr, et al., *How Americans Value Public Libraries in Their Communities* at <http://libraries.pewinternet.org/2013/12/11/section-1-an-overview-of-americans-public-library-use/>

⁹ U.S. CONST. Art I, Sec 8, Cl 8.

the exclusive rights of reproduction, distribution, making of derivative works, public display, and public performance.¹⁰ To guard public interest, though, explicit statutory exceptions were incorporated so that acts that would otherwise qualify as infringement would be permitted under certain circumstances. These exceptions range from service to the visually disabled, to compulsory licensing, to secondary transmissions by cable systems. The most generous of these exceptions are first sale and fair use which will be discussed below and are mainstays for library services.

Copyright is self-executing.¹¹ It requires only that an author create an original work and fix it in a tangible medium of expression.¹² Registration and notice are no longer required, though to file a civil suit for copyright infringement, the owner of a copyright must comply with certain legal requirements.¹³ For the purposes of this article, we will assume that any claims for infringement would be brought by a copyright owner who has fulfilled all statutory requirements for suit.

Direct infringement is generally viewed as a strict liability violation. Intent plays no part in the analysis, though intent and purpose can play a part in defenses and calculating damages. Central to most claims of infringement is the first exclusive right: reproduction. After all, a work often cannot be distributed, performed, or displayed without some sort of reproduction. There are only two elements to proving infringement of the reproduction right: ownership of the copyright by the plaintiff and copying by the defendant.¹⁴ The first, assuming that the author is also the copyright owner and has registered the copyright, is straightforward as copyright registration creates a presumption that the copyright is valid. The other requirement is copying, and this can be shown in several ways, but in the case of digitize-and-lend, no great imagination or lengthy explanation is required. Clearly, an entity that digitizes an item has actually copied the work. Given a strict reading of the right, a digitize-and-lend project would indeed infringe on the reproduction right of in-copyright works.

Even where infringement is proven or admitted, though, a defendant may assert an affirmative defense, one that shows that the act was legal even if infringement occurred. The ones most relevant to libraries are sections 107, 108, 109, and 121,¹⁵ and we will walk through the general role of each of these in libraries before turning to the application of some of them specifically to a digitize-and-lend endeavor.

¹⁰ 17 U.S.C. §106.

¹¹ If a book was authored before 1978, there were formalities required, so this statement represents only the law as it stands today for books authored after 1977. The inquiry into the validity of any copyrighted work created prior to that date is beyond the scope of this paper, but those interested are encouraged to consult NIMMER, *supra* note 3.

¹² 17 U.S.C. §102.

¹³ 17 U.S.C. §411.

¹⁴ See NIMMER, *supra* note 3, at §13.01.

¹⁵ 17 U.S.C. §§107, 108, 109, 121.

Section 121 permits the reproduction and distribution of published works to the visually impaired, so long as format used in reproduction is customized to meet the needs of that impaired group, the reproduction contains a notice that reproducing in any other format is prohibited, and the reproduction includes a copyright notice identifying the copyright owner and its date of original publication.¹⁶ One could argue that this type of use is now even more expansive, as courts have used fair use to supplement section 121, permitting distribution to the visually disabled even when the three statutory requirements are not met.¹⁷ The trend towards reducing requirements for providers to the print disabled may be due in part to the Marrakesh Treaty,¹⁸ which has 26 signatories including the United States and strongly supports copyright exceptions for the blind, visually impaired, and otherwise print disabled.

Section 108 authorizes libraries and archives to reproduce and distribute copyrighted works under certain conditions, including interlibrary loan, preservation, and replacement of obsolete formats. Each exception has a list of associated requirements and the section as a whole is read to recognize that libraries and archives serve a unique public purpose and therefore should have greater privileges than the general public in order to meet those obligations.

Section 109 is the codified expression of the first sale doctrine, which generally stands for the proposition that the legitimate owner of the material embodiment (e.g., book, DVD) of a copyrighted work may dispose of that embodiment in whatever way s/he sees fit.¹⁹ It is an explicit exception to the right normally reserved by authors “to distribute copies or phonorecords of the copyrighted work to the public by sale or other transfer of ownership, or by rental, lease, or lending.”²⁰ First sale is what allows libraries to lend works it has acquired.

Last, but in some ways, most important is fair use, which is codified in 107:

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—

¹⁶ 17 U.S.C. §121.

¹⁷ Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2014).

¹⁸ Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled, June 27, 2013, WIPO Doc. VIP/DC/8 Rev. at http://www.wipo.int/wipolex/en/treaties/text.jsp?file_id=301019

¹⁹ Bobbs-Merrill Co. v. Straus, 210 U.S. 339, 350 (1908).

²⁰ 17 U.S.C. §106(3).

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.²¹

Fair use is a broad exception, intended to be flexible enough to accommodate changes in customs and technologies, which makes it the only statutory provision that is not tied in some measure to the era in which it was drafted. It does not use language that assumes a certain type of technology nor does it require restrictions that may not make sense in different formats. It does not require that each factor be given a specific weight or even be seen as equal to each other. In fact, while the four enumerated factors are most commonly relied upon by courts, these factors are not exclusive,²² and courts appear to have considered other factors such as behavior of the parties,²³ free speech, and unexpected technologies.²⁴ Courts and Congress have stated it best:

The ultimate test of fair use... is whether the copyright law's goal of "promot[ing] the Progress of Science and useful Arts," U.S. Const., art. I, § 8, cl. 8, "would be better served by allowing the use than by preventing it."²⁵

[S]ince the doctrine is an equitable rule of reason, no generally applicable definition is possible, and each case raising the question must be decided on its own facts.²⁶

This paper will not explore sections 108, 109, or 121 in any greater detail, as those exceptions are each narrow in scope. While each can be used to justify some uses within a digitize-and-lend endeavor, only fair use can be used to justify the basic principle underlying the entire practice. To demonstrate how fair use applies, we will explore each component of a

²¹ 17 U.S.C. §107.

²² See NIMMER, *supra* note 3 at §§13.05[A][5][b] and [A][6].

²³ Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 549 (1985).

²⁴ Pamela Samuelson, *Unbundling Fair Uses*, 77 *FORDHAM L. REV.* 2537 (2009).

²⁵ Castle Rock Entm't, Inc. v. Carol Pub. Grp., Inc., 150 F.3d 132, 141 (1998).

²⁶ H.R. Rep. 94-1475 at 65 (1976).

digitize-and-lend strategy. The assumption in each case below is that some of a copyright owner's exclusive rights have been infringed and libraries would be asserting an affirmative defense of fair use in each case. Please note, however, that these are not the only facts or analyses possible, and in some cases, it may be possible to refute some of the charges of infringement entirely by using different theories or statutes. Since hypotheticals can vary widely, though, this paper seeks to use the likeliest scenario in its examination.

Digitization and Conversion

In order to create a digital collection that can be lent free from restrictions and tied to licensed e-works, a library must first undertake digitization. Therefore, this section explores digitization, separate from all other uses of the digitized item, as well as conversion when a technological format becomes disfavored, unstable or obsolete. Digitization itself implicates the reproduction right, as a print copy is scanned in and saved in a digital form. Similarly, conversion does the same, as it converts one digital format to another, creating a new copy in the process. Reproduction is not only one of the exclusive rights enumerated in 106,²⁷ but it is the only right that has remained protected throughout all of United States copyright history,²⁸ so has particular significance in copyright litigation.

There is a line of thought that digital copies cannot infringe on the reproduction right, as infringing the right requires a resulting "material object,"²⁹ but this is far from an uncontested view,³⁰ and it seems that the more logical and accepted view is that a digital reproduction is a reproduction that falls within the exclusive rights of a copyright owner. That is the view that this paper takes.

Not surprisingly, there have been no identified cases on these facts, probably because copyright owners have no way of knowing when these actions occur. Scrutiny is typically only triggered when a digitized item is used and that subsequent use comes to the attention of someone with a copyright interest in the work. Further, since the fair use evaluation requires use of some sort, it is difficult to apply in a reproduction-only case without additional uses.

²⁷ 17 U.S.C. §106(1).

²⁸ For the history of copyright and the adoption of the various exclusive rights in the colonies and the early United States, see Oren Bracha, *United States Copyright, 1672-1909*, in RESEARCH HANDBOOK ON THE HISTORY OF COPYRIGHT LAW (Isabella Alexander & H. Tomas Gomez-Arostegui eds, Elgar) (2016).

²⁹ See NIMMER, *supra* note 3, at §8.02[B][1]. See also the definition of copies and phonorecords in 17 USC §101.

³⁰ See *Capitol Records, LLC v. ReDigi Inc.*, 934 F. Supp. 2d 640, 649–50 (2013). ("Because the reproduction right is necessarily implicated when a copyrighted work is embodied in a new material object, and because digital music files must be embodied in a new material object following their transfer over the Internet, the Court determines that the embodiment of a digital music file on a new hard disk is a reproduction within the meaning of the Copyright Act.")

However, we do know that courts have regularly rejected the idea that the mere copying of the entirety of a work invalidates a fair use claim.³¹ In the context of libraries, digitization has largely turned into a non-issue, as evidenced by the recent *Google*³² and *HathiTrust*³³ opinions. Both of these cases stemmed from the Google Books project³⁴, which aimed to digitize and make available the world's literary history starting with the libraries of Harvard, the University of Michigan, the New York Public Library, Oxford, and Stanford. The number of participating libraries grew, even as authors and authors' groups filed lawsuits claiming infringement.

The Google Books project was scaled back considerably once litigation began, but enough of it survived for it to remain active in the courts. The corpus of digitized texts resided in multiple locations, but only two were the continued target of authors. One was Google itself, which made the texts searchable on their Google Books page, showing pieces of books where the search term was found. They also used the database they built for other purposes, such as creating the n-gram tool that could map the history of the use of a word or series of words. The second major target was HathiTrust, which used the works that Google had digitized as the central core of a shared digital repository. Unlike Google, they did not display any portions of the works, except to owning libraries or to the visually disabled, though they allowed users to search the database to determine what books in the collection contained those words.

In both of these cases, digitization itself was ignored in the court's final analysis or treated as incidental to other uses, not mentioned in either appellate opinion except in the descriptions of each project. There was no substantive evaluation of digitization's infringement of the reproduction right, despite the claim being a primary part of the initial complaint filed by the Authors' Guild against Google³⁵ and despite digitization itself being a non-transformative use.

In examining these two cases, it has become apparent that where downstream uses of the digitized works are determined to be fair, the digitization itself will also be considered fair, which requires us to look more closely at the subsequent uses for a work in a digitize-and-lend model.

Storage of Multiple Copies for Redundancy and Preservation

³¹ See *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 at 1165 (2007).

³² *Authors Guild v. Google, Inc.*, 804 F.3d 202, 229 (2015).

³³ *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 103–04 (2014).

³⁴ <https://www.google.com/intl/en/googlebooks/about/history.html>

³⁵ *Authors Guild et al v. Google Inc.*, Docket No. 1:05-cv-08136 (S.D.N.Y. Sept 20, 2005, complaint (filed 9/20/2005). (Third paragraph under Nature of the Action: "By reproducing for itself a copy of those works that are not in the public domain (the "Works"), Google is engaging in massive copyright infringement. It has infringed, and continues to infringe, the electronic rights of the copyright holders of those works.")

Once digitized, libraries must have a procedure for storage and retrieval of the resultant images. This type of use implicates the reproduction right, as best practices for the handling of digitized materials demand that the digital files meet minimum standards,³⁶ and that the images produced be preserved in multiple copies and stored in a secure manner while permitting checks for file fixity and data integrity.³⁷ The number of copies made should be stored in geographic locations with different disaster profiles to ensure that should one disaster in one area of a country destroy a copy that another will survive.³⁸ These extra copies are not intended for patron use but rather for preservation purposes and redundancy in the case of technology failure. Storage and access may also infringe on the public distribution right if the materials are accessed by users.

Both the *HathiTrust* and *Google* opinions largely sidestepped the issue of fair use as applied to storage, as the plaintiffs in the latter case could not establish any actual or realistic anticipated harm,³⁹ and in the former, lacked standing.⁴⁰ Some of the language in the *HathiTrust* case, though, clearly reflects the reasoning that would apply:

The Authors also contend that the copying is excessive because the HDL creates and maintains copies of the works at four different locations... But the record demonstrates that these copies are also reasonably necessary in order to facilitate the HDL's legitimate uses. Both servers contain copies of the digital works at issue. According to the HDL executive director, the “existence of a[n] [identical] mirror site allows for balancing the load of user web traffic to avoid overburdening a single site, and each site acts as a back-up of the HDL collection in the event that one site were to cease operation (for example, due to failure caused by a disaster, or even as a result of routine maintenance).” ...To further guard against the risk of data loss, the HDL stores copies of the works on two encrypted backup tapes, which are disconnected from the internet and are

³⁶ See Federal Agencies Initiative Still Image Working Group's Technical Guidelines for Digitizing Cultural Heritage Materials at http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf ; A Framework of Guidance for Building Good Digital Collections - NISO Recommended Practice at <http://www.niso.org/publications/rp/framework3.pdf> ; Technical Guidelines for Digitizing Archival Materials for Electronic Access <http://www.archives.gov/preservation/technical/guidelines.html>

³⁷ See the NDSA Levels of Digital Preservation: An Explanation and Uses at http://ndsa.org/documents/NDSA_Levels_Archiving_2013.pdf

³⁸ *Id.*

³⁹ *Authors Guild v. Google, Inc.*, 804 F.3d 202, 229 (2015) (“But on the present record, the possibility that libraries may misuse their digital copies is sheer speculation”)

⁴⁰ *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 103–04 (2014) (“The record... does not reflect whether the plaintiffs own copyrights in any works that would be effectively irreplaceable at a fair price by the Libraries and, thus, would be potentially subject to being copied by the Libraries in case of the loss or destruction of an original. The Authors are not entitled to make this argument on behalf of others, because § 501 of “the Copyright Act does not permit copyright holders to choose third parties to bring suits on their behalf.” *ABKCO Music*, 944 F.2d at 980”)

placed in separate secure locations on the University of Michigan campus. The HDL creates these backup tapes so that the data could be restored in “the event of a disaster causing large-scale data loss” to the primary and mirror servers.”⁴¹

From this language, we seek to do a full analysis of fair use for storage and preservation. The purpose and character of storage of multiple copies is both non-commercial and in the public interest, as each copy is not sold and the purpose is to ensure the preservation of an uncorrupted copy for current and future users. Further, preservation may be seen as a transformative use, a use that is “different in purpose, character, expression, meaning, and message from the page (and the book) from which it is drawn.”⁴² An author does not weave a whimsical tale to preserve it but rather to communicate the content. Transformative works receive special consideration in fair use determinations, as “the more transformative the new work, the less will be the significance of other factors ... that may weigh against a finding of fair use.”⁴³ With or without the transformative component, though, this factor should weigh in the favor of fair use because of society’s interest in preservation, even if copyright holders and courts might debate how many copies would be necessary to meet this purpose.

The second factor is the nature of the work, and in the case of digitize-and-lend, works of all types would be digitized. With fictional works, and unpublished works (e.g., manuscripts), as well as non-fiction works in the mix, it is possible that courts will view this factor as weighing against fair use. However, this factor has rarely been determinative in fair use cases,⁴⁴ and especially in the case of a transformative purpose, is viewed in light of the use of the work.⁴⁵ As in most cases, this factor would likely be non-determinative.

In weighing the third factor, the amount and substantiality of the work taken, the entire work would be taken for each title and the general principle is that the more of a work that is taken, the less likely that fair use will prevail.⁴⁶ However, courts have recently made clear that this prong of the analysis also cannot be viewed in isolation but must be viewed along with use.⁴⁷ As stated by the Court in Google:

Complete unchanged copying has repeatedly been found justified as fair use when the copying was reasonably appropriate to achieve the copier's transformative purpose and was done in such a manner that it did not offer a

⁴¹ *Id.* at 98-99.

⁴² *Id.* at 97.

⁴³ Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105 at 1111 (1990).

⁴⁴ PATRY ON FAIR USE, 4.1 (2015); *Authors Guild v. Google, Inc.*, 804 F.3d 202, 220 (2015).

⁴⁵ *Authors Guild v. Google, Inc.*, 804 F.3d 202, 219 (2015).

⁴⁶ See NIMMER, *supra* note 3 at §13.05(A)(3),

⁴⁷ See *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 96 (2014) and *Belmore v. City Pages, Inc.*, 880 F. Supp. 673, 678–679 (1995).

competing substitute for the original.⁴⁸

Lastly, as storage, even of multiple copies of a work, does not add copies to the market, there is no impact on the market and the copies do not offer “a competing substitute for the original.” Materials stored on secure servers are not generally accessible. Therefore, this last and most influential factor⁴⁹ would weigh in favor of fair use. At the end of the day, the author’s interest has not been affected and there is no market damage.

The transformative component of this argument would fail, though, if these copies are accessible by users, but that analysis will be reserved for the later section on lending.

Creation of a Full-Text Searchable Database

For resources to be identified as potentially useful by library patrons, libraries must make the materials findable. For digitized works, that can be done in several ways, the two most popular being a full-text searchable database and the addition of the database to a library discovery platform. Both accomplish the same goal, which is to allow the user to search the full-text of materials digitized, but the discovery platform goes one step further in intermingling the digitized content with other library resources (e.g., licensed content). Creation of a database would involve the reproduction right, as it requires digitizing all underlying works to populate the database.

Creating a searchable database has already been determined to be a transformative use by the court in *HathiTrust*,⁵⁰ as it serves the purpose of identifying materials relevant to a person’s research instead of serving as a substitute for the text of the books themselves. Adding such a database to a discovery platform serves the same purpose and therefore would fall under the same analysis.

As both uses are transformative and non-commercial, the purpose and character of use weigh in favor of a fair use determination. Per the *HathiTrust* court, the other factors are then viewed from the lens of a transformative work, and their weight lessens in the overall calculation. Each factor is considered in tandem with the purpose of the use. The analysis of the other factors would mirror the reasoning in the *HathiTrust* case. Factor two would be a neutral factor. While the works digitized would comprise both fiction and non-fiction, what matters in a transformative use is how the works are used, not how creative the works are.⁵¹

⁴⁸ Authors Guild v. Google, Inc., 804 F.3d 202, 221 (2015).

⁴⁹ Neil Weinstock Netanel, Making Sense of Fair Use, 15 Lewis & Clark L. Rev. 715 at 743-744 (2011). See also Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 549, 566 (1985) (describing the fourth factor as “undoubtedly the single most important element of fair use”)

⁵⁰ Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 97 (2014) (“we conclude that the creation of a full-text searchable database is a quintessentially transformative use”)

⁵¹ *Id.* at 98.

The third factor would come out in favor of fair use, as taking the whole of the works is necessary to satisfy the function of a full-text searchable database.⁵² And since the database does not substitute for the work, but merely as a finding aid, it has no market impact.⁵³

Displaying Snippets with Search Results

Context matters in a search, allowing users to assess whether a particular source is needed before retrieving or checking out the resource. Merely knowing that a term or phrase exists in a work is insufficient to narrow down potential research sources. Displaying snippets with search results implicates the reproduction, distribution, and public display rights and was most thoroughly reviewed in the *Google* case. As noted by that Court:

Google's division of the page into tiny snippets is designed to show the searcher just enough context surrounding the searched term to help her evaluate whether the book falls within the scope of her interest (without revealing so much as to threaten the author's copyright interests). Snippet view thus adds importantly to the highly transformative purpose of identifying books of interest to the searcher. With respect to the first factor test, it favors a finding of fair use (unless the value of its transformative purpose is overcome by its providing text in a manner that offers a competing substitute for Plaintiffs' books...⁵⁴

Factor two is non-dispositive in a transformative work; in evaluating the display of text snippets in the context for fair use, “[o]ne cannot assess whether the copying work has an objective that differs from the original without considering both works, and their respective objectives.”⁵⁵

Both factor three and factor four face a different test than that described for databases. In factor three, “[w]hat matters...is not so much “the amount and substantiality of the portion used” *in making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a public for which it may serve as a competing substitute.”⁵⁶ As showing limited snippets --- approximately one eighth of a page, in Google’s case--- cannot replace a book in full, limited displays would fall in a category that favors a finding of fair use.

Despite the court’s acknowledgement that snippets could result in the loss of sales, it

⁵² *Id.* at 98-99.

⁵³ *Id.* at 99. (“The Libraries contend that the full-text-search use poses no harm to any existing or potential traditional market and point to the fact that, in discovery, the Authors admitted that they were unable to identify “any specific, quantifiable past harm, or any documents relating to any such past harm,” resulting from any of the Libraries’ uses of their works (including full-text search))

⁵⁴ *Authors Guild v. Google, Inc.*, 804 F.3d 202, 218 (2015).

⁵⁵ *Id.* at 220.

⁵⁶ *Id.* at 222.

emphasized the longstanding principle that the loss of sales would “generally occur in relation to interests that are not protected by the copyright,”⁵⁷ such as the discovery of a fact (e.g., president’s birthday) that could be conveyed entirely in a snippet. The court went even further in noting that while it was possible for snippets to contain enough words to encroach on protected expression, the brevity and disconnected nature of snippets could not substitute for the entire work.⁵⁸ On that premise, the fourth factor favored fair use.

The *Google* court concluded: “The purpose of the copying is highly transformative, the public display of text is limited, and the revelations do not provide a significant market substitute for the protected aspects of the originals.”⁵⁹ Since library discovery engines also only display small snippets, we believe that the same analysis in the Google Books case would apply to the discovery of resources in digitize-and-lend projects.

Controlled Circulation of Digitized Materials

By far and away, the most challenging component in digitize-and-lend projects is the lending of titles in a manner which substitutes for the original. Lending here would be controlled by some mechanism that has digital rights management protocols (e.g., Overdrive), to limit simultaneous access and to avoid duplication or further distribution by users. The mechanism would also enable the use of specific checkout periods or recall.

Lending of digitized items potentially involves the reproduction, public distribution and derivative work rights. Lending of a digital item would be impossible without digitization, which would be a reproduction normally protected by 106. Similarly, user access to the full-text of the works would meet the description of the public distribution right within the same statute. Whether the right to make derivative works would be infringed is a more difficult matter.

The text of the derivative works right is contained in 106(6), and the boundaries of the right may prove to be troublesome to define in a technological age. Whether or not the derivative work right has been infringed will depend in part on the tool used by a library to disseminate e-books and in part on courts’ understanding that technology can automate instantaneously actions once seen to infringe this right. For example, screen readers essentially make audio versions from any text, and various applications can create translations from machine-readable text. Both of these products are identified as paradigmatic examples of a derivative work in the definition section of the Copyright Act.⁶⁰ It may be possible for libraries

⁵⁷ *Id.* at 224.

⁵⁸ *Id.*

⁵⁹ *Id.* at 229.

⁶⁰ 17 U.S.C. §101. (“A ‘derivative work’ is a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted. A work

to infringe on this right should they create a reader that has within it a translation or audio function. Then again, if those functions do not result in a new embodiment (e.g., new digital file), these functions would not meet the permanence required for an infringing act. As it is highly unlikely that a library would create its own reader but instead would use or recommend an existing product on the market, for the purposes of this discussion, it is assumed that libraries would not infringe on this right.

The character and purpose of use is identical to the use of the original, so lending does not appear to be a transformative use, though an argument could be made that it is in cases where use is intended for disabled patrons or those in remote locations. In those instances, the digital work is being used primarily to overcome distance, not to replicate content. However, since the lower court in *HathiTrust* essentially made this assertion for the provision of texts to those with visual impairments and was soundly rebuked by the appellate court,⁶¹ we will use the latter's more limited interpretation of transformative use.⁶² The digitized version of the book is being used in the exact same manner as the original print version --- to read the content --- and therefore cannot be transformative. The character of use is non-commercial. At best, the combination of those elements would cause this factor to be seen as neutral since the overall purpose of lending aligns with the general non-commercial, public interest mission of a library to provide resources to its community. At worst, it would be seen as weighing against fair use.⁶³ Given how the *Google* and *HathiTrust* courts treated digitization, though, which is also a non-transformative use, it is reasonable to take the view that this factor is more neutral than negative, and a fair use claim can survive despite the failure to produce a non-transformative use.

As in all cases, factors two and three cannot be viewed in isolation. It is true that the works lent comprise both fiction and non-fiction, and that the full-text of the works have been copied. However, it would also be true that the amount taken is that which is necessary to the use, as reading a book loses its value if you do not have access to the entirety of the book. Given the general consensus that these factors cannot be viewed in isolation and are themselves not determinative, these factors would seem neutral.

consisting of editorial revisions, annotations, elaborations, or other modifications which, as a whole, represent an original work of authorship, is a 'derivative work'.")

⁶¹ *Authors Guild, Inc. v. HathiTrust*, 902 F.Supp.2d 445, at 461 (2012). (holding that providing copies to the print disabled was a transformative use) and *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 101 (2014) (correcting the district court, and determining that normally the copies provided to print disabled would be considered derivative works).

⁶² *Id.* at 101.

⁶³ *Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1262 (2014) ("...Defendants' use of excerpts of Plaintiffs' works is not transformative. The excerpts of Plaintiffs' works ...are verbatim copies of portions of the original books which have merely been converted into a digital format. Although a professor may arrange these excerpts into a particular order or combination for use in a college course, this does not imbue the excerpts themselves with any more than a *de minimis* amount of new meaning.")

Fair use, then, seems to rest primarily on the fourth factor, as it does in most cases. Since this is not a transformative work, the question is: does the lending of a digitized copy of a work affect “the potential market for or value of **the copyrighted work?**” [emphasis added] The key part of this test is the definition of the market. It is the market of the copyrighted work that is relevant, not any other market. In the case of digitize-and-lend, this becomes complicated as two (or more) markets may be intertwined: the market for the copyrighted work and the market of a given technology. After all, if a library buys a copy of a work in print, digital, and audio formats, are each of these purchases being made because of the content in each? Or is it possible that some of these purchases are made for the format / technology?

Let us walk through three examples to untangle the markets involved:

- Library One has purchased one copy of a title, it format shifts but continues to use the original format as well, thereby turning one copy of a work into two copies. In this case, the library has not paid for two copies, but only one, and is using two to meet user demand.
- Library Two has purchased one copy of a title, it digitizes the title and only uses the digitized copy. It places the original print copy into a dark archive, not accessible to the public. It has paid for one copy and only uses one copy.
- Library Three shows a hybrid situation. A library has purchased one copy of a title, it digitizes the title, and primarily uses the digitized copy. The print copy is in a staff-only area, and not available for general browsing or circulation. On occasion, a user specifically requests the print copy. In those instances, the library takes the digital copy out of circulation, and checks out the print copy. While this library has two copies, only one copy is used at any given time.

Library One clearly could not avail itself of a fair use defense as it is using its digitized copy in a manner that increases the number of copies of the work available for its users. That additional copy substitutes for a copy that should have been purchased for its content and therefore actually affects the market for the work. In the second two instances, the effect is distinctly different. The digitized copy does not replace or substitute for a copy available on the free market, but rather, it replaces *the library's legitimately acquired copy*. Libraries Two and Three are not looking to gain more copies than they have legitimately acquired, but simply want to make full use of the content that they have.

Even though Libraries Two and Three have not increased the number of copies in use, it is still possible that their actions impact the market for the works digitized. But as noted in

Campbell, not every action that lessens or kills demand for the original produces a harm recognized by copyright laws.⁶⁴ The most common examples are parody or blistering criticism, both of which could reduce sales of a work. Copyright does not protect the work against these types of market effects.

Similar arguments can be made here. First, copyright is medium neutral,⁶⁵ a concept which has been used both offensively and defensively in litigation. An author of a work cannot claim a new copyright if all he does is transfer an existing work to a new medium,⁶⁶ and a defendant cannot fight a charge of infringement by claiming that converting a work to a new format removes the work from the protection of the original copyright.⁶⁷ In other words, copyright protects the **work**, not a format.

Market impact, then, should be measured on the overall market for the work, not by any change in any of the individual markets for the different formats of the work. In a digitize-and-lend strategy, libraries would still buy titles in print representing the actual number of copies needed by their communities. Libraries may also continue to buy new formats of the same work if those new formats contain new content desired by the community, but where they merely need access to the content of the item they already own, format shifting allows them to continue to use the content of the investments they have already made. At its heart, the question that has to be asked and answered is this: if a library has one lendable copy before digitization and one lendable copy afterwards, can market harm (if any) be attributable at all to infringing the work? Or is it entirely attributable to technology, which copyright does not protect?

If a library decides to purchase materials only in print, this may indeed reduce overall library spending on the same titles in other formats. But the library could also decide to purchase more copies in print than they do now to meet community demand, knowing that they will have the ability to format shift, and that demand is sufficient to require multiple copies; such action would cause an uptick in print purchases to counteract the decline in digital purchases of the same title. Even if the rebalancing of sales does not occur and copyright owners see an overall decrease in the number of copies of a title sold to libraries, that decrease may not be associated with the market for the copyrighted work.

⁶⁴ *Campbell v. Acuff-Rose Music, Inc.* 510 U.S. 569, 590-91 (1994).

⁶⁵ *New York Times v. Tasini*, 533 US 483 at 502 (2001). (“the transfer of a work between media’ does not ‘alte[r] the character of that work for copyright purposes”).

⁶⁶ *L. Batlin & Son, Inc.*, 536 F.2d at 491. (“[A] copyright claimant’s production of a work of art in a different medium cannot by itself constitute the originality required for copyright protection.”)

⁶⁷ In the early 1900s, there were a series of cases where defendants, creators of perforated piano rolls, escaped liability for reproduction because the technology they used was novel. See as an example, *White-Smith Music Pub. Co. v. Apollo*, 139 F. 427 (1905). The current medium neutrality language in section 102 --- “Copyright protection subsists... in original works of authorship fixed in any tangible medium of expression, **now known or later developed...**” (emphasis added) --- was a deliberate addition to avoid future inequities. See H.R.Rep. No. 94-1476, at 52 (1976), reprinted in 1976 U.S.C.C.A.N. 5659, 5665.

In such a decline, a strong argument can be made that the profit lost would be sales to libraries that used to buy new formats for the technology, not the content. After all, if a library consistently purchased copies of a work in multiple formats before format shifting (e.g., 1 print, 1 digital), but then became perfectly satisfied with a single copy once it could digitize and lend (see Library Two above as an example), that means that the needs of its community are satisfied by a single copy. The extra purchase pre-format-shifting was not driven by the need for an extra copy of the content but by the technological container of the content.

Since potential purchases lost in this scenario are caused by a demand for newer technology and not additional copies of the same content, one can argue that the market for the copyrighted work was temporarily and artificially elevated by non-copyright interests. Any reduction in profits caused by libraries format-shifting would actually correct that artificial market. Loss of a technology-created windfall to authors and publishers should not be part of the fair use analysis, as the loss would not be associated with the aspect of the market attributable to the substantive, creative works protected by copyright.

Second, it is also possible that library users might purchase fewer copies themselves, choosing to rely more on libraries as their titles become available online. Any reduction in sales by users is highly speculative, though, as there may be no way to prove that those users would ever have wanted the books enough to purchase them. A reduction could be caused by several factors, from ease of use (e.g., library digital copy has a built-in screen reader) to readers finding library resources more readily because of the enhanced search function of databases. Declining sales for these reasons should not be within the fair use analysis, as they have to do more with usability and access than the making available of an added, unauthorized copy of the content of the work being read or checked out.

Further, the fact that these actions would be taken by non-profit libraries for non-commercial purposes requires copyright owners to meet a higher burden of proof when it comes to the fourth factor. "If the intended use is for commercial gain, that likelihood [for market harm] may be presumed. But if it is for a noncommercial purpose, the likelihood must be demonstrated." ⁶⁸ Even in *Google*, where the defendant was a commercial entity, the court ruled for fair use where the plaintiffs failed to furnish sufficient evidence to support a reasonable theory of harm in potential hacking of servers to gain access to digitized copies. The bar in digitize-and-lend, where libraries would be defendants, should be at least that high, if not higher.

Third, even if a copyright owner could produce data demonstrating direct correlation between library lending specific copies and an overall reduction in sales in the corresponding titles, they would be unable to prove that this reduction was due to an added copy to the market. Libraries are permitted to lend their legitimately acquired collections under 109. As long as they circulate only the number of copies they have legitimately acquired, any market

⁶⁸ Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 451 (1984).

harm in circulating a digital copy instead of a print one logically would be tied to the technology itself (e.g., faster delivery) more than the content. This is not a market harm protected against by copyright.

In short, though we have a non-transformative use, the purpose remains library lending of content that a library legitimately owns,⁶⁹ and any shifts in market behaviors would be attributable to technology, not the “market for or value of the copyrighted work.” Controlled circulation of digitized works should therefore be fair use.

Use of Digitized Materials By Another Library to Replace Print Equivalent

Like the lending structure, this issue has not yet been reviewed or decided by the courts. In these situations, a library could obtain from another library a copy of a digitized text owned by each library, which would implicate the reproduction and distribution rights. Both the providing and receiving libraries would then have a digital copy for use. Neither library would seek to save the costs of acquiring the work itself, as they would each need to independently own legitimately acquired copies. Each would control circulation of any digital items lent, and at no time would they circulate more copies than what they own. The receiving library would seek to reduce the costs to itself of digitization, but if this is seen as a market harm, it would be a market harm to the industries creating digitization equipment, not to the owners of copyright. Any reduction of sales incident to format shifting from print to digital would fall under the analysis in the section immediately preceding this one, leading us to the same conclusion: this type of use would be fair.

Resource Sharing through Interlibrary Loan or Consortia

Libraries often share resources with one another through ILL or consortial arrangements⁷⁰ and digital items simplify many of the processes for sharing. The fair use analysis for this type of use within digitize-and-lend models mirrors that for controlled lending of items except for the potential market effect.

Practically speaking, all of these uses are already permitted in an analog world. Libraries ship materials to each other through ILL, libraries with multiple branches will purchase materials for the whole instead of replicating every collection in each branch, and consortia or partner libraries will enter into agreements to allocate shared specialty purchases to different libraries.

⁶⁹ In this context, it is important to note that owning content is not the same as owning copyright. Owning the content means legitimately acquiring an embodiment of the work which, by definition, includes one copy of the content.

⁷⁰ Examples include: <https://www.btaa.org/projects/library/reciprocal-borrowing/uborrow>, <https://www.railslibraries.info/services/interlibrary-loan>, <https://www.orbiscascade.org/shared-content-program/>

What changes in a digital realm is speed and costs (staff, materials, and shipping), so there may well be a market effect for any given title. It is theoretically possible for libraries collectively to buy fewer copies of a given title if that title is shared by several libraries, freeing up resources to acquire more unique content or more copies of titles in greater demand. It is equally plausible that use of digital items would result in greater demand. As libraries within consortia would still need to acquire the number of copies necessary to sustain its community, the digitize-and-lend model could result in a net gain of purchases, with all operational savings being reinvested into collections. As either of these outcomes is highly speculative, and the only data available on print v. digital versions of titles is highly variable,⁷¹ this is an area where hard data would be necessary to the analysis.

The fair use argument survives even without this data, though, as any market effect is again caused by non-copyrighted-content elements. It is not the content that would drive cost savings or makes it easier to collaborate; it is the technology. More efficient use of technology -- such as FedEx or other shipping methods in the past --- can have real market effects, but these market effects are not harms that copyright laws protect against.

Big Data and Data Mining

While not an explicit goal of digitize-and-lend projects, whenever a substantial data store is created, there is the possibility and promise of data mining. Any large collection of information can be used to identify patterns in a way that the use of any single document cannot. Whether through creation of tools like Google's n-gram⁷² or through circulation/search analysis, libraries with such databases can gain information that otherwise would be invisible. Even while protecting users' privacy, they would be able to see what titles and subjects are most popular at any given time, evaluate types of use (reference v. reading) based on the amount of time users spend with works, and how social media (or any media) effects the

⁷¹ David J. Gray & Andrea J. Copeland, *E-Book versus Print: A Per Title Cost and Use Comparison of a Public Library's Popular Titles*, 51 REF. & USER SVCS Q 334 (2012) at <https://journals.ala.org/rusq/article/download/3915/4384> (print equivalents circulate more than digital versions); Justin Littman & Lynn Silipigni, *A Circulation Analysis of Print Books and E-Books in an Academic Research Library*, 48 LIB. RESOURCES & TECH SVCS 256 (2004) at <http://www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/lrts/archive/48n4.pdf> (ebooks get 11% more use than print counterparts); Deborah M. Taylor, *Comparison of Selected e-Books and Equivalent Print Books: Have Handheld Portable Devices Increased Use in Three Aggregated Resources?*, 10 J. ELEC. RESOURCES IN MED. LIBR. 11 (2013) (ebooks are used exponentially more than print equivalents); Cathy Goodwin, *The e-Duke Scholarly Collection: e-book v. print use*, 33 COLLECTION BUILDING 101 (2014) (e-books show greater interest but less substantive use).

⁷² <https://books.google.com/ngrams> See also the Court's mention of the n-gram tool in *Authors Guild v. Google, Inc.*, 804 F.3d 202, 217 (2d Cir. 2015) ("...through the ngrams tool, Google allows readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods. We have no doubt that the purpose of this copying is the sort of transformative purpose described in *Campbell* as strongly favoring satisfaction of the first factor.").

interest in subjects or titles. These are uses where the full-text of any text need never be viewed or could just be viewed in snippet form.

While not fully explored by the courts, the analysis here should closely mirror that of the searchable database. The purpose of the use would be transformative, seeking to use the full-text of books not for the substantive content but rather to determine historical patterns, identify relationships between works, possibly uncover authors of anonymous or pseudonymous works through language patterns, or identify misprintings through the comparison of multiple copies of a text. Once the transformative test is passed, the rest of the factors fall into place. The nature of the works would include both non-fiction and highly creative fictional works, but given the use, this would simply result in a conclusion that the second factor is non-determinative. As the full-text of each work is needed to create a database to be mined in such a fashion, the full reproduction of works is reasonably necessary to accomplish the use. And since the use cannot substitute for the original works, there is no market damage. Data mining, so long as the full-text of works is not displayed, would therefore be fair use.

Other Common Questions and Their Answers

1. Doesn't the fourth factor in fair use require a broader examination of market effects, as articulated in *Campbell*?

The *Campbell* court noted that the fourth factor of fair use "requires courts to consider not only the extent of market harm caused by the particular actions of the alleged infringer," but also "whether unrestricted and widespread conduct of the sort engaged in by the defendant ... would result in a substantially adverse impact on the potential market" for the original.⁷³ While this is indeed a broader test, the framework contained in this paper applies only to libraries and is based on their unique status within society and copyright law. It cannot and would not apply generally to the public or to commercial entities, neither of which would have the same non-profit, society-focused purpose of a library or the interests in controlling circulation.

2. Are there other factors a Court can or should consider for fair use?

Since the factors in section 107 are not exclusive, courts are welcome to consider other factors. Since case law provides few examples of these, they are not as easy to define or examine as the traditional prongs, but scholars have posited a few additional factors that could be helpful to libraries. The first is a focus on uses that promote learning⁷⁴ as a possible

⁷³ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994) (Citing Nimmer § 13.05[A][4]).

⁷⁴ Samuelson, *supra* note 24 at 2580-87.

separate factor favoring fair use, an approach that would benefit academic libraries in particular as their collections, and the deployment of them, are tied to the educational mission of their parent institutions. The second calls for equity to be its own factor, where courts would essentially ask which is more fair or reasonable: (1) to ask libraries to expend “scarce resources on the same material multiple times because of technological changes,”⁷⁵ or (2) to allow them to format shift to use content purchased. If they are able to lend materials on the same terms as they lend the physical equivalents, are the authors harmed such that the fair use defense should be denied?

3. Does the *ReDigi* case have any impact on this analysis?

Observers of digitize-and-lend projects have wondered about the impact on digitize-and-lend based on two cases that sought to undertake similar projects with respect to digital music. Both failed, though for reasons that are not entirely extendable to library digitize-and-lend efforts. Those two cases are *ReDigi*⁷⁶ and *MP3.COM*⁷⁷.

ReDigi is a case about a “used music” service that allows a user to sell a legitimately acquired digital music file to another user, deleting it from the devices of the first before allowing the second to gain access. *MP3.COM* provided a service where subscribers could store, customize and listen to the music contained on their CDs from any place where they had an Internet connection. Subscribers would provide “proof” of ownership by inserting their CDs into their computers and allowing *MP3.COM* to verify the CD before providing the user access to the music online. At the time that *MP3.COM* provided its services, music companies were not yet selling/licensing their music in online sound files. Since the issues raised in both cases were largely the same, this analysis will treat them as one and will use *ReDigi* as the representative of both, as its court opinion contained more substance than *MP3.COM*’s.

The fair use analysis in both of these cases was minimal, and both concluded that the combination of a non-transformative use, commercial purpose, and likely market damage doomed the fair use claim.⁷⁸ As market damage may be presumed in commercial uses,⁷⁹ the courts in each case were highly critical of the plaintiffs’ fair use defenses. Libraries using digitize-and-lend have a non-commercial purpose, and as noted in an earlier section, damage cannot be presumed in such instances. Since the first and fourth factors play the largest roles in fair use, and a library’s purpose is central to the analysis for both factors, the conclusions in *MP3.COM* and *ReDigi* are inapplicable here.

⁷⁵ Michelle M. Wu, *Collaborative Academic Library Digital Collections Post- Cambridge University Press, HaithiTrust and Google Decisions on Fair Use*, 1 J. COPYRIGHT IN EDUC. & LIBRARIANSHIP 1, 17-18 (2016).

⁷⁶ *Capitol Records, LLC v. ReDigi, Inc*, 934 F.Supp.2d 640 (2013).

⁷⁷ *UMG Recordings v. MP3.COM*, 92 F.Supp.2d 349 (2000).

⁷⁸ *Id.* at 351-52; *Capitol Records, LLC v. ReDigi, Inc*, 934 F.Supp.2d 640, 653-54 (2013).

⁷⁹ *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 451 (1984).

The first sale defense was eliminated summarily as well, as it applies only to the disposal of a legitimately acquired copy of an embodiment of a work. In cases of digitize-and-lend, a library is not arguing to lend the original physical item, which certainly would be protected by first sale, but rather a proxy for that item. First sale, therefore, is not directly relevant to digitize-and-lend.

However, the *ReDigi* case remains important, not for its decision but for its analysis, which could be misinterpreted in a manner beyond the limited context of the decision. The first potential issue is its statement that infringement of the reproduction right rests on whether a new copy is made instead of an additional copy and infringement of the distribution right on the sharing of that new copy.⁸⁰ This statement should be properly limited to the examination of the elements of infringement and should not play a role in evaluation of affirmative defenses such as fair use, where an additional copy added to the market should be evaluated differently from a replacement copy that does not increase the numbers of works in circulation, or in other cases where a copy is legally made (e.g., under section 108). Otherwise, libraries would be faced with the incongruous result of being permitted to make a digital copy to replace damaged items but being unable to convey it to another library or archive should the first library close or otherwise be unable to maintain it.

Second is simply a typographic error that carries with it significant consequences if taken at face value. That typographic error can be found in this statement “...the **fair use** defense is, by its own terms, limited to the assertions of the *distribution right*.”⁸¹ [boldface added] If fair use were to be limited only to the distribution right, then many of the earlier cases on fair use, including *Sony*, would be undermined, as those cases dealt strictly with the reproduction right. Since the *ReDigi* court immediately follows this puzzling statement with a citation to 109, though, it should be apparent to all readers and to other courts that the distribution limitation was intended to apply to first sale and not fair use.

Third is the court’s view of the newly created digital file as an unlawful reproduction, which could give rise to fears that anyone creating a new digital file from a copyrighted work would be seen as making an unlawful reproduction. However, context matters in this case. The file produced in *ReDigi* was unlawful because the court determined that fair use did not apply. If the court had found a fair use, then the copy would be a lawful one, and the rest of the court’s reasoning would not follow. This is not a far-fetched concept, as academic and educational enterprises have received special dispensation by the courts for acts that would be unfair for commercial enterprises.⁸²

⁸⁰ Capitol Records, LLC v. ReDigi, Inc, 934 F.Supp.2d 640, 648-49 (2013).

⁸¹ *Id.* at 655.

⁸² Cambridge Univ. Press v. Patton, 769 F.3d 1232, 1266 (2014) (“Consistent with these principles, courts will not sustain a claimed defense of fair use when the secondary use can fairly be characterized as a form of “commercial exploitation,” i.e., when the copier directly and exclusively acquires conspicuous financial rewards from its use of

Last, in reviewing whether or not first sale applied in this instance, the court engaged in a lengthy discussion of digital works and how they could not be treated the same as physical embodiments. The opinion relied heavily on the DMCA Report⁸³ and implied that copyright provided greater protections for works in digital format than for those in print format. Again, though, in reading the opinion in full, this reading is confined only to first sale, not to digital items in other contexts. While one could also challenge the assumptions by the *ReDigi* court and the DMCA report by noting that copyright does not protect format, it protects the work, and at some point in time, technology may well be able to completely transfer one digital item from one location to another; the validity of digitize-and-lend efforts does not require this step. Digitize-and-lend relies on fair use, not first sale.

4. What if a digitized version of a title owned by a library is available for sale by the publisher?

The fair use defense for digitize-and-lend does not depend on the absence of a digital copy for sale. Practically speaking, though, there are very few vendors who sell digital works, at least in the academic market. The vast majority of works are licensed, which by definition conveys fewer rights than purchase. If they were available for purchase for a fair price, libraries may well purchase them in that format, simply to avoid the costs and hassle of digitization, but they should not be obligated to do so. A library that can more effectively use its resources by digitizing content it already owns should be permitted to do so even if an identical copy is available for sale by the publisher. After all, the library has already paid once for the content; it is not reasonable to require it to pay again.

5. Does *Texaco*⁸⁴ impact the analysis at all?

Texaco involved a library at a for-profit organization, where the library routed journal issues to its researchers. If researchers expressed an interest in having a copy of an article in these issues, the library would provide a photocopy of it. Considering the circumstances and the number of copies made, the court determined that this use was not fair use. *Williams* is a case with facts almost identical to *Texaco*'s but the actor was a non-commercial, research library.⁸⁵ The same action determined to be infringement in *Texaco* was determined to be a fair use in *Williams*. *Texaco* itself is not applicable to library digitization-and-lend projects, but the

the copyrighted material. Conversely, courts are more willing to find a secondary use fair when it produces a value that benefits the broader public interest.”).

⁸³ DMCA Section 104 report (August 2011).

⁸⁴ American Geophysical Union v. Texaco, 60 F.3d 913 (1994).

⁸⁵Williams & Wilkins, Co. v. U.S., 203 Ct.Cl. 74 (1973).

difference in treatment between the two cases should support the argument that non-commercial library activity should be evaluated differently than commercial activity.

6. What about international copyright and use by users in other countries?

Since controlled circulation and other technological mechanisms (e.g., geofencing) would permit libraries to restrict access only to certain patrons and/or geographic areas, they need not tackle this particular issue immediately. This is a fortunate circumstance, considering the different perspectives of nations on intellectual property.

Despite the Berne Convention⁸⁶ and other major copyright treaties, copyright laws around the world vary widely, which is why this paper does not spend time discussing use in other locations. Very few nations recognize fair use in the way that the United States does, and issues such as the public lending right, compulsory royalties, and moral rights change the calculation for each jurisdiction in such a way that it would be impossible to cover in a single paper.

It should be noted, though, that while the analysis of this paper has been confined to the United States, other countries (e.g., Canada, Israel) have been moving closer to a copyright regime that respects uses such as the ones contemplated by fair use in the United States.

Conclusion

Technology cannot be acceptable only when it shifts the balance towards the interests of copyright holders, as online markets, advances in transportation, or advances in printing did. Each of these lowered the cost of doing business and allowed publishers to reach a broader range of buyers. At the end of the day, any changes in the market in digitize-and-lend models are not caused by flooding the market with unauthorized copies of a copyrighted work. They are changed by a technological solution to a service problem that simply was not present at an earlier point in our history.

Libraries have been slow to adopt digitize-and-lend due to issues ranging from equipment costs to debates over minimum standards for digital images to a fear of copyright laws. This article seeks to address the concerns about copyright and redirect energies to practical solutions that libraries face every day on providing our users with access to our collections.

⁸⁶ Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, 828 U.N.T.S. 221 (revised at Paris July 24, 1971).