



Georgetown University Law Center
Scholarship @ GEORGETOWN LAW

2016

Collaborative Academic Library Digital Collections Post-Cambridge University Press, HathiTrust and Google Decisions on 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/1797>
<http://ssrn.com/abstract=2838898>

J. Copyright Educ. Libr. (2016)

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>



Part of the [Intellectual Property Law Commons](#), and the [Law and Society Commons](#)

Collaborative Academic Library Digital Collections Post- Cambridge University Press, HaithiTrust and
Google Decisions on Fair Use
Michelle M. Wu¹

Academic libraries face numerous stressors as they seek to meet the needs of their users through technological advances while adhering to copyright laws. This paper seeks to explore one specific proposal to balance these interests, the impact of recent decisions on its viability, and the copyright challenges that remain after these decisions.

The challenges facing academic law libraries are many, but the three primary ones are budget, demand, and misperceptions. Though actual means and medians of collection expenditures continue to grow,² they have failed to keep pace with inflation rates,³ resulting in a net decrease in spending power over the last decade. On a different front, student and faculty appetites for multiple formats and interdisciplinary research sources continue to expand, placing greater strain on shrinking budgets. Exacerbating the effect of both of these is the provision of resources through digital means, resulting in the invisibility of the library and the common misperception that libraries are no longer necessary.

Considering this landscape, it may seem odd to propose digitization as a potential solution, as this step would make even more resources available online and heighten the division between a library and its resources. However, user expectations and habits have made clear that online access is more heavily relied upon than other resources,⁴ and libraries need to meet users where they are to remain relevant. This article does not seek to resolve the long-standing tension between use and funding, but instead aims to provide some relief for the budgetary and demand issues.

The proposal in question was described at an earlier stage in the *Law Library Journal*,⁵ but has evolved to one less dependent on the forming of a consortium. At its heart, the proposal is that academic law libraries digitize their holdings, share them with one other through a controlled-circulation mechanism, and leverage their financial resources more effectively through a collaborative collection development and maintenance function. The first piece is the simplest of the whole, asking participating libraries to

¹ Michelle M. Wu is the Associate Dean for Library Services and Professor of Law at Georgetown University Law Center.

² Charles B. Lowry, ARL Budgets After the Great Recession, 2011-13 at <http://publications.arl.org/rli282/2> and ARL Law Library Statistics at <http://www.arlstatistics.org/about/series/law>

³ "Prices of U.S. and Foreign Published Materials."

http://www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/collect/serials/ppi/LMPI_2014Article.pdf For historical data, see <http://www.ala.org/alcts/resources/collect/serials/spi>

⁴ Michelle M. Wu and Leslie A. Lee, *An Empirical Study on the Research and Critical Evaluation Skills of Law Students*, 31 Legal Reference Services Q. 205-238 (2012).

⁵ Michelle M. Wu, Building a Collaborative Digital Collection: A Necessary Evolution in Libraries, 103 L. Lib. J. 527 (2011).

digitize their materials, prioritizing materials that have no online equivalent (primarily monographs). The second piece is more challenging, as it requires libraries sharing these digitized resources to do several things: remove the print title digitized from circulation and enter its digitized version into a centralized repository. The repository would allow circulation of any digital item, but would restrict the number of simultaneous users to the number of copies “held” by the repository (i.e., if partner libraries own ten copies of a title, the repository would have the authorization to circulate ten copies of the digital version at a time, so long as all ten print equivalents are not being circulated). Any circulation of an item would be controlled through an online lending platform like Overdrive⁶ or Open Library⁷, and any digital object would be further contained through digital rights management (DRM), limiting the ability to make copies, print, or loan the item to someone else. In other words, the lending mechanism would duplicate the existing circulation and interlibrary loan functions of a library, just in digital form. The proposal itself is extendible to other types of libraries or digitization projects, but as each type of library or project has unique characteristics, the analysis here will focus on academic law libraries as an example.

Such a solution contemplates on-going reduction of costs in several ways. Sharing of physical materials is costly, once one aggregates the costs of shipping both ways, the personnel required to pull and ship materials, and the time lost in shipping. The sharing of digital materials reduces or eliminates these costs. There would no longer be any shipping costs, with circulation accomplished online, nor would library personnel need to retrieve or send materials, as all of this could be automated. Such an approach also reduces the likelihood of loss of materials or the time investment in negotiating with recalcitrant patrons to return items; online lending platforms enable immediate “reclaiming” of an item.⁸ On the flip side, of course, are the costs of digitization itself, but as only one library would need to digitize a title held by many, the costs are less daunting when shared by all partners. Further, the practical reality is that while funders may be reluctant to fund print acquisitions, there are many more funding options for digitization.

There would also be a reduction of cost to the researcher. Presently, the novice researcher often only searches for online materials, not realizing the treasure trove of resources available only in print. Even though discovery platforms allow the simultaneous searching of print and e-resources, the fact that one set of resources is full-text while the other only contains basic bibliographic information constrains the search. The digitization of printed materials ensures a level playing field for all resources, regardless of their original formats. All could be searched simultaneously and in an equivalent fashion. Even the expert scholar could see savings in the time necessary to identify the titles necessary to her research and in retrieving the item. There would be even greater savings in the use of special collections, where the researcher often has to travel to the owning library’s location to access the resource. If these collections digitized and made available through an online platform, the time and expense of travel may be reduced or eliminated.

⁶ <https://www.overdrive.com/>. Description of other platforms can be found in John Novak, eBook Lending Platforms, 25 *Against the Grain* 22-26 (2013).

⁷ <https://openlibrary.org/>

⁸ Technologically, this often isn’t a reclaiming but a control mechanism to restrict access to a title. However, for the majority of the user population, it functions the same way.

Though many libraries have expressed interest over the years in such a project, almost every conversation on the topic has stopped once in-copyright materials come under discussion. The three exceptions are where litigation is unlikely or where there is an accepted exception: orphan works, providing digital materials for disabled persons, and limited-access archiving. It is not that librarians doubt that building and using such a collection is fair use, but their (and their Universities') anticipation of the threats of litigation and the associated costs have stunted academic library exploration into more broadly useful digital collections.

The recent decisions in *Cambridge University Press v. Patton*, *Authors Guild, Inc. v. HathiTrust*, and *Authors Guild, Inc. v. Google* all breathe life into efforts to build a working collaborative digital library, removing some perceived barriers, and allowing libraries to concentrate on narrower copyright issues. This article will provide readers with a brief review copyright and a description of notable eras within fair use for libraries, before advancing to a discussion of the most recent court opinions and their spawning of a new era.

Part I: Copyright and Fair Use

Article I, Section 8, Clause 8 of the U.S. Constitution reads:

Congress shall have the power...[t]o promote the progress of Science and the Useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.

Known as the Copyright Clause, this declaration shows that copyright was always a means to an end, not an end in and of itself. Granting rights to authors to protect their works was seen as necessary to ensure continued creative output,⁹ and despite research showing that neither remuneration nor a long term of protection is necessary for rapid and robust growth of the creative commons,¹⁰ this general principle remains the motivating force for many copyright laws.

In furtherance of the Copyright Clause, then, Congress undertook the drafting of several copyright acts, the most extensive of which is the one still in effect today, albeit in somewhat altered form: the Copyright Act of 1976. The act's framework is largely structured to favor the copyright owner, and the broadest of its provisions can be found in the granting of exclusive rights to copyright owners for limited times, captured in in section 106 and 106A. These authors' rights are then followed by (mostly) narrow exclusions in sections 107 through 122. Section 107 is notable in these exclusions, as its protections are not narrowly circumscribed. It is drafted in terms as broad as the original grant. Section 107 speaks to

⁹ Melville B. Nimmer and David Nimmer, *Nimmer on Copyright* §1.03 (Matthew Bender, Rev. Ed.)

¹⁰ Rebecca Tushnet, *Economies of Desire: Fair Use and Marketplace Assumptions*, 51 *William & Mary L. Rev.* 513 at 513 (2009) ("...the desire to create can be excessive, beyond rationality, and free from the need for economic incentive. Psychological and sociological concepts can do more to explain creative impulses than classical economics.")

fair use, an affirmative defense to copyright infringement, permitting uses of copyrighted works in conditions that advance societal interests and do not rob the author of the fruits of his labor. Section 107 reads:

- (a) 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—
 - (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.
- (b) 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.

This language in the first part of the statute is the codification of the common law concept of fair use as articulated by Justice Joseph Story in *Folsom v. Marsh*.¹¹ Though Justice Story set forth the factors, he did not provide any definitions or weights to any of the factors. As he himself said,

This is one of those...questions, arising in the administration of civil justice, in which it is not, from the peculiar nature and character of the controversy, easy to arrive at any satisfactory conclusion, or to lay down any general principles applicable to all cases...[I]n cases of copyright, it is often exceedingly obvious, that the whole substance of one work has been copied from another, with slight omissions and formal differences only, which can be treated in no other way than as studied evasions; whereas, in other cases, the identity of the two works in substance, and the question of piracy, often depend upon a nice balance of the comparative use made in one of the materials of the other; the nature, extent, and value of the materials thus used; the objects of each work; and the degree to which each writer may be fairly presumed to have resorted to the same common sources of information, or to have exercised the same common diligence in the selection and arrangement of the materials. Thus, for example, no one can doubt that a reviewer may fairly cite largely from the original work, if his design be really and truly to use the passages for the purposes of fair and reasonable criticism. On the other hand, it is as clear, that if he thus cites the most important parts of the work, with a view, not to criticise, but to supersede the use of the original work, and substitute the review for it, such a use will be

¹¹ *Folsom v. Marsh*, 9 F. Cas. 342 (D. Mass. 1841).

deemed in law a piracy. A wide interval might, of course, exist between these two extremes, calling for great caution and involving great difficulty, where the court is approaching the dividing middle line which separates the one from the other.¹²

In other words, while he could articulate what could be considered, he felt that the very nature of intellectual property was too protean to allow for the easy application of rules. Other courts seemed to agree, as between 1841 and 1976, they often cited *Folsom* but did not follow a consistent formula in application. Congress agreed that flexibility was needed and sought to retain it when it codified the concept of fair use.¹³ In crafting the statutory language, though, Congress inadvertently made an already complex concept even more difficult to understand or apply. First, it delineated a four factor test that, while non-exclusive, was easiest to apply if treated as all-inclusive. As law is an institution that depends in part upon consistency and precedent, courts are not well equipped to deal with a statute that both prescribes a test but provides no instructions for application. Second, it included an exemplar of fair use --- “teaching (including multiple copies for classroom use)” --- that could be read as an exception to the four factors. Third, it included in its legislative history the Agreement on Guidelines for Classroom Copying in Not-For-Profit education Institutions with Respect to Books and Periodicals.¹⁴ Though this agreement explicitly opens with

The purpose of the following guidelines is to state the minimum and not the maximum standards of educational fair use under Section 107... The parties agree that the conditions determining the extent of permissible copying for educational purposes may change in the future; that certain types of copying permitted under these guidelines may not be permissible in the future; and conversely that in the future other types of copying not permitted under these guidelines may be permissible under revised guidelines.

the fact that it contains prohibitions on conduct contradicts the language about minimum standards. As a whole, then, the codification of fair use served to provide no more clarity than Justice Story’s longer description of fair use, and in fact, made fair use more difficult to apply than in earlier times.

The statute is often described as being in three parts: the preamble including exemplars of fair use; the (non-exclusive) four-pronged test of fair use; and a statement about unpublished works (added in 1992 to counter the effects of *Harper & Row*). Of these, the second part --- the four pronged test --- has received the most attention, both by judges and by scholars. Though there is disagreement as to whether there really are four prongs, or only one factor with multiple facets,¹⁵ courts continue to address all four points as distinct ones in their analyses.

The four factor test, which had been intended to provide a minimum floor for the issues to be considered in determining if a use was a fair one, has taken on an entirely different meaning over the

¹² Id. at 344-45.

¹³ H.R. Rep. No. 94-1476 (1976)

¹⁴ Id.

¹⁵ Barton Beebe, An Empirical Study of U.S. Copyright Fair Use Opinions, 1978-2005, 156 U. Penn. L. Rev. 549 at 617 (2008).

years. It has morphed from a fuzzy test, much like what Justice Story had described, to a rigid application of the four factors, and only the four factors.¹⁶ The first factor, nature and character of use, was relevant insofar as it spoke to the two interests of copyright. Where use was commercial, there was a greater likelihood that the use would not be fair. Copyright laws were intended to support an author's or copyright owner's ability to reap the rewards of her own work or investment, and commercial use of a copyrighted work by someone other than the owner, therefore, ran contrary to this intent. If the nature and character of use, on the other hand, were informative (e.g., news reporting) and non-commercial, this factor was more likely to weigh towards fair use. This factor is one of the most influential factors, and in recent years has become the most influential factor,¹⁷ for reasons that will be detailed in later sections. Note that "commercial" is not the same as for-profit, and there are instances where an entity can receive direct or indirect rewards from use of a copyrighted work while still being seen as non-commercial in nature.¹⁸

The second factor, the nature of the copyrighted work, recognizes a sliding scale of protection applying to copyrightable works dependent on their content. As copyright rewards creativity, more creative works are seen as more deserving of copyright protection than others. This concept is supported in copyright even beyond the examination of fair use. For example, there are explicit exclusions within copyright laws and regulations for categories of works that are not protected (e.g., facts),¹⁹ and cases have repeatedly limited copyright protection for works that draw heavily on fact or common knowledge.²⁰ In fair use assessments, "[the] law generally recognizes a greater need to disseminate factual works than works of fiction or fantasy."²¹ Therefore, the less creative the work, the less the use will weigh against fair use. This factor is one of the easiest to evaluate, but is typically not definitive.²²

Amount and substantiality is the third factor, and at one time, was as simple to weigh as the second factor. The more of a work that was taken, the more likely the factor would weigh against fair use, and there was a time when the taking of an entire work ended the analysis.²³ Since the advent of duplicating technologies, though, courts have had to adapt their analyses. Where copying an entire work in an analog world would have been infringement, the caching of entire works by a computer (e.g., website) to assist in transmission is not seen as infringing on the rights of the copyright owner,²⁴ whether under the theory that the reproduction is temporary or under the fair use defense. In fair use, amount and substantiality has become a factor that cannot be evaluated independently. Instead, it merges with

¹⁶ Id. At 561-62.

¹⁷ Neil Weinstock Netanel, *Making Sense of Fair Use*, 15 Lewis & Clark L. Rev. 715 at 743-744 (2011).

¹⁸ *Cambridge University Press v. Patton*, 769 F.3d 1232 at 1264-1267 (2014).

¹⁹ 17 U.S.C. §102(b); 37 C.F.R. §202.1.

²⁰ *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340 (1990) (limiting protection of basic directories); *Baker v. Selden*, 101 U.S. 99 (1879) (limiting protection of forms); *Hoehling v. University City Studios, Inc.*, 618 F.2d 972 (noting that scenes a faire are not protected by copyright).

²¹ *Harper and Row*, 471 U.S. 539 at 563.

²² *Supra* note 15 at 610.

²³ *Worldwide Church of God v. Philadelphia Church of God, Inc.*, 227 F.3d 1110 at 1118 (2000) ("copying an entire work militates against a finding of fair use.")

²⁴ William F. Patry, *Patry on Copyright* §21:39

factor one so that the test is now whether “the amount and substantiality of the portion used ... are reasonable in relation to the purpose of the copying.”²⁵ Since this factor depends on another, it also is rarely definitive in resolving a dispute where fair use is raised.

The last factor, market effect, has always been the most difficult to identify and evaluate. While this factor is still evolving, courts have generally agreed on some basic guidelines. Nimmer describes the test as “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.²⁶ Where the defendant’s work substitutes for the copyright owner’s, this factor will weigh heavily against fair use. Market effects that are indirect --- such as reducing the market for an author’s work by making a popular, stinging parody of it --- generally will not be counted as a market effect for the purposes of this analysis. Up until the last decade, this was the most influential factor and predicted in over 95% of the cases the outcome of any fair use analysis.²⁷

The factors, at first glance, seem straightforward even if broad, but the courts, in their interpretation, have tended to confuse more than clarify. In fact, some scholars have criticized the courts heavily for seeming to bend the prongs of fair use to whatever ends they seek.

Courts tend first to make a judgment that the ultimate disposition is fair use or unfair use, and then align the four factors to fit that result as best they can. At base, therefore, the four factors fail to drive the analysis, but rather serve as convenient pegs on which to hang antecedent conclusions.²⁸

Though empirical studies have shown that courts have been less arbitrary than Nimmer suggests,²⁹ they have also demonstrated why there continues to be considerable confusion over fair use. First, the Supreme Court has declined to explicitly correct prior interpretations of fair use, even where lower courts have split in their application of Supreme Court precedent. The clearest example of this is in *Sony*, where the Court had stated in dicta, “every commercial use of copyrighted material is presumptively an unfair exploitation of the monopoly privilege that belongs to the owner of the copyright.”³⁰ Lower courts cite this language repeatedly, and the Supreme Court has had multiple opportunities to correct the misapplication of *Sony*’s statement. Instead, it has muddled the waters further by saying both that the application was wrong but then reiterating the statement that caused the confusion in the first place.³¹

²⁵ *Campbell v. Acuff-Rose*, 510 U.S. at 586–587 (1994).

²⁶ *Supra* note 9 at § 13.05[A] [4]

²⁷ *Supra* note 15 at 617.

²⁸ David Nimmer, “Fairest of them All” and Other Fairy Tales of Fair Use, 66 *Law & Contemp Probs.* Winter-Spring 2003 at 281

²⁹ *Supra* note 17 at 721.

³⁰ *Sony Corp. of America v. Universal Studios, Inc.*, 464 U.S. 417 at 451 (1984).

³¹ *Supra* note 15 at 599-601, describing the Supreme Court’s contradictory statements on commercial effect and fair use from *Sony* through *Campbell*.

With this historical background, libraries' and universities' fears of litigation are unsurprising. Pursuing a course of digitization when litigation is likely and the outcome unpredictable increases the potential loss to the entity undertaking the action. Not only could universities be held liable for damages for infringement plus attorneys' fees (though these might be mitigated or remitted to \$0 by 17 U.S.C. 504(c)(2)), but if found to have infringed, they would lose access to the digitized works as well their investment of time and effort to create those electronic documents. Without certainty, libraries have no safe way forward.

Before reviewing the various fair use eras, it is important to note that two other statutes are relevant when talking about libraries and fair use, as they intersect regularly with 107 analyses. These are sections 121 (17 USC 121) and section 108 (17 USC 108). In brief, section 121 recognizes how difficult it may be for visually impaired individuals to obtain copies of reading materials in usable formats. Therefore, section 121 carves out an exception to entities that provide specialized formats for the blind or disabled. Section 108 carves out copyright exceptions, including preservation and interlibrary loan, for libraries and archives under certain conditions so long as the entities strictly comply with limits placed on the type of reproduction that can be done. While neither of these sections plays a significant role in this article, both are necessary to the operations of libraries as they navigate copyright and fair use.

The next section will explore the development of fair use and libraries over four different eras.

Part II: Fair use eras

When Congress drafted the Copyright Act of 1976, the specter of technology was already looming. The most contentious sections were those where technology was seen to be potential game changers, and fair use was one of these sections.³² Photocopiers had been introduced to businesses in the late 1950s and were becoming common in libraries and archives in addition to businesses. The final language of the act, along with its legislative history, reflect the drafters' struggle with technology, especially in light of the fact that it was so new that actual impact could not be predicted. Little could they have foreseen exactly what an impact technology would make to copyright and fair use.

To better highlight technologically driven advances in fair use, this paper summarizes major developments by dividing case law up by eras. For simplicity's sake, most eras have been named below for the case in the era that sparked the shift that defines that time period.

Pre-Technology Era (pre-1983)

Though photocopiers and other technologies (e.g., recording devices) did exist in this era, they were all analog and largely burdensome. For that reason, they are included in the pre-technology stage, despite their advances over more manual duplication efforts. In this era, the analysis of the four factors in any

³² William F. Patry, *The Fair Use Privilege in Copyright Law* at 351 (2d ed, 1995).

case was casual at best. While there was a slight shift in the approach in 1978, when the current Copyright Act (and section 107) went into effect, fair use was a rarely asserted defense. Libraries did not face much scrutiny, and only one relevant case was found.

*Williams & Wilkins*³³ in 1973 involved a publisher suing the library of the National Institutes of Health for photocopying articles for their employees. The library routed a copy of each journal to those interested, and upon request, the library would photocopy an article for the requesting researcher. It would not make multiple copies for any researcher, typically limited copies to only one article from any given journal issue for any one researcher, and limited the number of pages that it would copy in any request. The library neither monitored the reasons for the requests nor did they require that the materials be returned. The court found the use to be fair, as the purpose of the copying was solely for the development and dissemination of knowledge, the works copied were factual in nature, the library had established reasonable limits on how much could be copied, and there was limited market effect.

The decision was split four to three, appealed to the Supreme Court, and was affirmed due to an equally divided Court. The Copyright Act of 1976 was passed after the *Williams* case, and two pieces within its legislative history make it appear as if the case had had some influence on the development of fair use and Congress' view of it. The first is the House Report accompanying the legislation, which outlines the relationship between 107 and 108 for libraries:

The Register of Copyrights has recommended that the committee report describe the relationship between this section and the provisions of section 108 relating to reproduction by libraries and archives. The doctrine of fair use applies to library photocopying, and nothing contained in section 108 "in any way affects the right of fair use." No provision of section 108 is intended to take away any rights existing under the fair use doctrine. To the contrary, section 108 authorizes certain photocopying practices which may not qualify as a fair use.³⁴

This serves to reiterate *Williams*' recognition of the unique status of libraries and their importance to society. The second, less deferential provision, is from the Senate Report on 108:

Subsection (g)... does not authorize the related or concerted reproduction of multiple copies of the same material whether made on one occasion or over a period of time, and whether intended for aggregate use by one individual or for separate use by the individual members of a group. For example, if a college professor instructs his class to read an article from a copyrighted journal, the school library would not be permitted, under subsection (g), to reproduce copies of the article for the members of the class.

While it is not possible to formulate specific definitions of "systematic copying," the following examples serve to illustrate some of the copying prohibited by subsection (g).

³³ *Williams & Wilkins Co. v. U.S.*, 487 F.2d 1345 (1973).

³⁴ *Supra* note 13 at 74.

1. A library with a collection of journals in biology informs other libraries with similar collections that it will maintain and build its own collection and will make copies of articles from these journals available to them and their patrons on request. Accordingly, the other libraries discontinue or refrain from purchasing subscriptions to these journals and fulfill their patrons' requests for articles by obtaining photocopies from the source library.
2. A research center employing a number of scientists and technicians subscribes to one or two copies of needed periodicals. By reproducing photocopies of articles the center is able to make the material in these periodicals available to its staff in the same manner which otherwise would have required multiple subscriptions.
3. Several branches of a library system agree that one branch will subscribe to particular journals in lieu of each branch purchasing its own subscriptions, and the one subscribing branch will reproduce copies of articles from the publication for users of the other branches.³⁵

Though legislative history is not binding, the examples provided by the Senate were illuminating in light of the Williams case, as it signaled that the use that had been deemed fair in that case perhaps should not have been. Even as the Senate made these determinations, it recognized that libraries were evolving and that Congress would need more guidance on these issues. Congress subsequently established the National Commission on New Technological Uses of Copyrighted Works (CONTU) in an attempt to bring greater resolution to technology and copyright.³⁶ Unfortunately, CONTU encountered the same difficulties that Congress had. Technologies were still rapidly changing, so much so that it was impossible to come to agreement on terms that would survive their evolution.

The House and Senate Reports also provide libraries with clues on how Congress viewed libraries and copyright in this era. Congress used language in 108 that intentionally broadened protections for libraries beyond those available for other entities. The language adopted also almost exclusively dealt with one of the six rights --- reproduction. Clearly, Congress was primarily concerned about the commercial impact of reproduction as opposed to rights related to distribution or derivative works, issues that would later become as important.

Sony (1984-1993)

In 1984, *Sony* dramatically changed the view of fair use. In this case, the court had to determine whether Sony's marketing and selling of the Betamax recorder, a device designed to duplicate copyrighted works (i.e., television programs), was copyright infringement.³⁷ The *Sony* court made three valuable contributions to fair use in its decision, at least in relation to libraries and their uses of technology. First, as noted in the description of the factors above, wholesale copying of a work was

³⁵ S. Rep. No. 94-473 (1976) at pages 70-71

³⁶ Public Law 93-573

³⁷ Supra note 30.

once presumptively unfair. Sony was the case that modified this factor, acknowledging that copying the whole of a work could be fair under certain circumstances.³⁸ Second was its recognition that technologies used to infringe can also have substantial non-infringing uses, and that removing the technology from the market because of infringing uses could inflict great societal harm.³⁹ Balancing these interests, the Court permitted the technology (i.e., Betamax recorder) to continue to be distributed and set the precedent for newer technologies to receive the same treatment. Last, the *Sony* court made it more difficult to prevail in copyright infringement cases where the use was non-commercial by stating that “[a] challenge to a noncommercial use of a copyrighted work requires proof either that the particular use is harmful, or that if it should become widespread, it would adversely affect the potential market for the copyrighted work.”⁴⁰ This is a higher burden than mere substitution, as the copies made in this case clearly could substitute for the original.

Thanks to the Sony holding, companies developing and improving equipment for mass reproduction – like high-speed digitization equipment --- prospered.

Campbell (1994-2013)

While the term “transformative use” was coined before *Campbell*,⁴¹ *Campbell* marked the point at which transformative use became a common element in the analysis of fair use’s factor one.⁴² The court in *Campbell* determined that 2 Live Crew’s parody of Roy Orbison’s “Pretty Woman” was fair use despite actual copying and a commercial purpose. The heart of the court’s analysis rested on “whether the new work merely supersedes the objects of the original creation, or whether and to what extent it is “transformative,” altering the original with new expression, meaning, or message. 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.”⁴³

While the facts of *Campbell* do not generally apply to libraries, the concept of transformative use and its influence on fair use analyses are incredibly important not only to libraries but to all users in an era where technological advances made it easier to copy, manipulate, and use copyrighted works in unusual or unanticipated ways. The graphical web came into being in this time frame, as did Google and other major search engines, spawning a series of cases that would test the limits of fair use as applied to innovative technologies.

³⁸ Supra note 30 at 450 (“Moreover, when one considers the nature of a televised copyrighted audiovisual work, see 17 U.S.C. § 107(2), and that timeshifting merely enables a viewer to see such a work which he had been invited to witness in its entirety free of charge, the fact that the entire work is reproduced, see *id.*, at § 107(3), does not have its ordinary effect of militating against a finding of fair use.”)

³⁹ *Id.* at 456

⁴⁰ *Id.* at 451

⁴¹ *Twin Peaks Production Inc. v. Publications Intern, Inc.*, 996 F.2d 1366 (1993); *Basic Books, Inc. v. Kinko’s Graphics Corp.*, 758 F. Supp. 1522 (1991).

⁴² Supra note 17 at 737.

⁴³ Supra note 25 at 569.

The most notable of these were *Arriba Soft*⁴⁴ and *Perfect 10*,⁴⁵ both of which involved search engines which crawled the web and displayed thumbnails of images on other websites in their search results. In *Arriba Soft*, the program downloaded the full-sized picture, generated smaller thumbnails to display in search results, and then deleted the full-sized pictures from their servers. Clicking on any of the thumbnails would in-line link to the full-sized image. Google Images' caching of webpages and search of images was the dispute in *Perfect 10*. The plaintiffs in both cases alleged copyright infringement of copyright owners' reproduction, display, and/or distribution rights. Since the purpose of both services was to improve access to information not to create artistic expressions, most of the uses in both cases were determined to be transformative and fair.

These cases combined established a foundation on which libraries and other entities built larger digitization projects. The courts recognized that copyrighted works combined could serve needs beyond those met by each individual work and that such a combination could produce a transformative work benefitting the public and deserving of special consideration in copyright infringement cases. How entities used these concepts to digitize and make available library collections is covered in the next section.

Before moving to that section, though, it is necessary to note that despite the steps forward in the evolution of fair use doctrine, this era also dealt a setback to libraries in the form of a 1994 case similar to *Williams*, but with a notably different outcome. Texaco employed hundreds of researchers and its library ran a routing and photocopying service nearly identical to the one run by NIH in *Williams*.⁴⁶ The primary difference between the two cases was purpose, with NIH performing research services for the government and Texaco researching to improve its commercial performance in the petroleum industry. Texaco was found to be liable for copyright infringement, causing some concern and confusion among libraries. In both *Williams* and *Texaco*, library staff provided copies of journal articles for researchers. The only distinction was in the nature of each entity's business. Given how divided the *Williams* court was and the subsequent Senate Report with language on what should not count as fair use, the Texaco decision made some libraries question if their routing and copying practices were protected or not.

HathiTrust (2014-)

A series of recent cases builds on the cases in earlier years to make up our present fair use era.

The first of these cases was *Authors Guild v. HathiTrust*. HathiTrust is an organization formed by the libraries that participated in the Google Books project and was founded with the purpose of preserving library materials. Participating libraries deposit their works with HathiTrust, which indexes and stores the works within its repository. It allows all users to search the materials stored, but only owning libraries to view the full-text. Search results note in which works, on what pages, and with what

⁴⁴ Kelly v. Arriba Soft Corp., 336 F.3d 811 (2002).

⁴⁵ Perfect 10, Inc. v. Amazon.com, 508 F.3d 1146 (2007).

⁴⁶ American Geophysical Union v. Texaco, 60 F.3d 913 (1994).

frequency the search term(s) are located, but do not display the full-text of the items themselves. Plaintiffs filed suit against HathiTrust and its member libraries, claiming copyright infringement in the use and storage of materials. The court determined that the digitization and storage of copyrighted works was fair use, as was the use of the digitized works in a database, so long as the full-text of the works remained unseen by the public. Since HathiTrust's database did not supply full-text access to works to anyone other than those supported by section 121, its services were not seen to substitute for original works in a way that would create market harm.

For digitization efforts, this ruling brought greater clarity on permitted uses of copyrighted works. The *HathiTrust* court had pronounced that "the creation of a full-text searchable database is a quintessentially transformative use,"⁴⁷ and that in such cases, the other factors in the fair use inquiry would be given less weight. The court had also recognized that in order to create a fully searchable database, it was necessary for the libraries to make copies of the full texts of the works, and that this wholesale copying was permitted under fair use.⁴⁸ And, finally, the last part of the holding repudiated of the authors' claim that storage of these works in multiple sites was prohibited by copyright.⁴⁹ Libraries' reproduction and storage of multiple copies served efficiency and preservation interests and therefore were not seen as infringing.

The second case in the series was *Cambridge University Press*, in which Georgia State University operated two services where digitized or e-copies of works were posted online for student use. One was an e-reserves system where materials were uploaded by libraries, and the second was a course management system where faculty uploaded works themselves. In both cases, access to the works were restricted to the students enrolled in the respective courses and works could not be uploaded until faculty had completed a fair use analysis. While this court remanded the fair use analysis to the lower court, it did make several pronouncements that are relevant to library digitization processes. The first was resolving the seeming conflict between the *Williams* and *Texaco* cases. The court here explicitly noted that nonprofit educational use may be granted greater protections than the same actions by for-profit entities, even when the educational institution might derive some indirect profits from their actions.⁵⁰ The second was a firm reiteration that the four factors are neither exclusive⁵¹ nor intended to be applied rigidly.⁵²

The last decision in the series was the Second Circuit's in relation to the Google Books Project. As Google displays more of a work than HathiTrust does, though still not full-text except with permission of

⁴⁷ *Authors Guild v HathiTrust*, 755 F.3d 87 at 97 (2014).

⁴⁸ *Id.* at 98 ("Because it was reasonably necessary for the HDL to make use of the entirety of the works in order to enable the full-text search function, we do not believe the copying was excessive")

⁴⁹ *Id.* at 99 ("We have no reason to think that these copies are excessive or unreasonable in relation to the purposes identified by the Libraries and permitted by the law of copyright. In sum, even viewing the evidence in the light most favorable to the Authors, the record demonstrates that these copies are reasonably necessary to facilitate the services HDL provides to the public and to mitigate the risk of disaster or data loss.")

⁵⁰ *Cambridge University Press v. Patton*, 769 F.3d 1232 at 1263-1267 (2014).

⁵¹ *Id.* At 1282.

⁵² *Id.* At 1258-1260.

the copyright owner, this case exerted greater pressure on the boundaries of fair use. The authors raised arguments that had not appeared within HathiTrust, including claims that enabling search infringed on authors' derivative rights, that snippets actually could serve as substitutes for original works, and that Google's distribution of copies of digitized works to contributing libraries was infringing. The court summarily dismissed the plaintiff's claim to have a derivative right in the search function as "an author's derivative rights do not include an exclusive right to supply information (of the sort provided by Google) about her works."⁵³ It also gave short shrift to the argument that Google's snippet view substitutes for the original work, as there was no series of searches that would result in the entirety of a book being displayed. Even though plaintiffs were able to demonstrate that they could view a fair amount of a book with some effort, the number of hours required to do this was seen to be so unusual as not to serve as a realistic portrayal of likely user behavior.

The court found snippet views to be transformative as they provide context about a word that cannot be gained by the word alone⁵⁴ and that "at least as presently structured by Google, the snippet view does not reveal matter that offers the marketplace a significantly competing substitute for the copyrighted work."⁵⁵ Further, despite recognizing that plaintiffs might indeed lose sales due to Google's snippet views was insufficient to meet the test for substitution.⁵⁶ Also important to digitization projects was the recognition that libraries can use contractors or other entities to digitize their collection.

The contract between Google and each of the participating libraries commits the library to use its digital copy only in a manner consistent with the copyright law, and to take precautions. In these circumstances, Google's creation for each library of a digital copy of that library's already owned book in order to permit that library to make fair use through provision of digital searches is not an infringement. If the library had created its own digital copy to enable its provision of fair use digital searches, the making of the digital copy would not have been infringement. Nor does it become an infringement because, instead of making its own digital copy, the library contracted with Google that Google would use its expertise and resources to make the digital conversion for the library's benefit.⁵⁷

Remaining copyright challenges for library digitization

Returning now to the proposal to build a collaborative academic law library collection, we examine the challenges remaining even after *Google*. As a reminder, the proposal is to digitize collections to level the

⁵³ Authors Guild v. Google, Inc., 804 F.3d 202 at 207-208 (2015).

⁵⁴ Id. at 218. ("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.")

⁵⁵ Id. at 222

⁵⁶ Id. at 224. ("But the possibility, or even the probability or certainty, of some loss of sales does not suffice to make the copy an effectively competing substitute that would tilt the weighty fourth factor in favor of the rights holder in the original. There must be a meaningful or significant effect "upon the potential market for or value of the copyrighted work.")

⁵⁷ Id. At 229.

playing field for users searching for information and to more effectively collaborate on collection development, access, and maintenance.

The legitimacy of the creation, storage, and use of searchable database from copyrighted materials is no longer in question, as both HathiTrust and Google have affirmed these uses as fair. Similarly, the issues surrounding the actual digitization process have been resolved. Therefore, only one major challenge to fair use remains. Throughout all case law, one principle remains intact, that a copy that supplants or substitutes for the original for the same purpose is not fair use.⁵⁸ The one notable exception is in *Sony*, where copies of television programming made by private citizens in their own homes with recording technology was not considered to be infringing. However, the analysis was limited solely to the recording function, and not the performance or distribution to others,⁵⁹ so it has limited application in library digitization projects that include the goal of facilitating the lending of library materials.

As we saw above, there have been cases where wholesale copying has occurred and been declared fair use --- *ArribaSoft* and *Perfect 10* being examples --- but in those cases, the digital originals were available freely on other sites, the copies were seen to be inferior to the originals, and therefore not substitutes. The proposal for a digital collaborative collection differs in that the originals are not freely available online, and the copies should not be inferior in quality to the originals, unless digital is seen to be inferior to print.

This history and consistent messaging, though, should not discourage libraries from undertaking this project and risking litigation. The reason this principle remains unchallenged is because no project has accomplished circulation of in-copyright digitized documents in a manner that comports with the spirit of copyright. The purpose and manner of use proposed for this collaborative collection differs from any other case heretofore before the courts, and therefore, it presents an issue of first impression on which libraries can press for a less cited but more appropriate test for fair use: “the use must be of a character

⁵⁸ Folsom at 344–345 (“[A] reviewer may fairly cite largely from the original work, if his design be really and truly to use the passages for the purposes of fair and reasonable criticism. On the other hand, it is as clear, that if he thus cites the most important parts of the work, with a view, not to criticise, but to supersede the use of the original work, and substitute the review for it, such a use will be deemed in law a piracy.”); *Harper & Row* at 568 (fourth factor concerned with “use that supplants any part of the *normal* market for a copyrighted work”) (emphasis added); *HathiTrust* at 95 (“A fair use must not excessively damage the market for the original by providing the public with a substitute for that original work.”); *Google* at 221, 223 (“The larger the amount, or the more important the part, of the original that is copied, the greater the likelihood that the secondary work might serve as an effectively competing substitute for the original, and might therefore diminish the original rights holder’s sales and profits” and “on whether the copy brings to the marketplace a competing substitute for the original, or its derivative, so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may opt to acquire the copy in preference to the original”); *Williams* at 1366 (“It is undisputed that the photocopies in issue here were exact duplicates of the original articles; they were intended to be substitutes for and they served the same purpose as the original articles. They were copies of complete copyrighted works within the meaning of Sections 3 and 5 of the Copyright Act. This is the very essence of wholesale copying and, without more, defeats the defense of fair use.”)

⁵⁹ *Supra* note 30 at 425.

that services the copyright objective of stimulating productive thought and public instruction without excessively diminishing the incentives for creativity.”⁶⁰

The current four prongs for fair use analysis have increasingly shown their age and inability to adapt to technology, and it often feels as if courts are obligated to go through the analysis even when they know that the only elements are the first and fourth factors. The second and third prong are largely irrelevant in cases of digitization, with almost all cases turning on the transformative nature of the use and commercial impact. In discussion of transformative use, Courts struggle to distinguish between transformative uses justifying a fair use defense and the right to make derivative works, which belongs exclusively to the copyright owner. Recent cases have shown how challenging this division is, never quite coming up with a satisfying or clear distinction. The courts in *HathiTrust* and *Google* both made the attempt, but because both of their premises are easily overcome, demonstrate their ineffectiveness.

The court in *HathiTrust* based its transformative assessment on a belief that “[t]here is no evidence that the Authors write with the purpose of enabling text searches of their books.”⁶¹ However, there are books that are indexed in detail. If authors could prove that they did write with such an intent, but just executed it poorly, would the analysis of transformative use have changed? Many authors write books without intending them to be adapted into movies, and yet when they are so adapted, that adaptation is considered a derivative right.

The court in *Google* picked up where *HathiTrust* left off, basing its reasoning on what it felt were the differences between transformative uses qualifying as fair use and infringement of derivative rights:

A further complication that can result from oversimplified reliance on whether the copying involves transformation is that the word “transform” also plays a role in defining “derivative works,” over which the original rights holder retains exclusive control...The statute defines derivative works largely by example, rather than explanation. The examples include “translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgement, condensation,” to which list the statute adds “any other form in which a work may be ... *transformed*.” ..As we noted in *Authors Guild, Inc. v. HathiTrust*, “[p]aradigmatic examples of derivative works include the translation of a novel into another language, the adaptation of a novel into a movie or play, or the recasting of a novel as an e-book or an audiobook.” ... While such changes can be described as transformations, they do not involve the kind of transformative purpose that favors a fair use finding. The statutory definition suggests that derivative works generally involve transformations in the nature of *changes of form*”⁶²

⁶⁰ Pierre N. Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105 at 1110 (1990).

⁶¹ *Supra* note 47 at 97.

⁶² *Supra* note 53 at 215-216

The court then continues with the seemingly devastating statement, “If Plaintiffs’ claim were based on Google’s converting their books into a digitized form and making that digitized version accessible to the public, their claim would be strong.”⁶³ Taken to its conclusion, building a collaborative digital library would be doomed. Not only is it a change in form, which falls squarely into the court’s definition of a derivative work, but if libraries were to lend materials to all of their users, the public could very well be part of that group. I do not think that this is what was intended by the court.

By resetting the fair use test to the principles in the Copyright Clause instead of using the section 107 test, courts would be free to advance the two principles that animated copyright in a meaningful way even in an age of rapid technological change. Courts should be able to recognize that a shift in form and making such a form available to others even beyond the §121 exception can serve a public good and can be legitimate in a world where technology changes daily. Failure to do so creates a windfall for authors that was never intended at the time of the nation’s founding. Copyright laws assume that an economic incentive is needed for creation of new works, but that economic incentive should not be multiplied simply by the introduction of new technology. Otherwise, each time a technology dies and is replaced by another, the author could sell the same work to the same buyer. That may produce a greater economic incentive to create, but it also causes large negative market effects in that users would need to invest more funds in the same work instead of spreading that investment over multiple works. Ultimately, this approach would undermine the societal benefit interests inherent to copyright.

Our hypothetical digital library would only permit circulation of copies equal to those that had been purchased by the participating libraries. Circulating a digitized version of a book does not expand a library’s authority beyond the uses contemplated at initial purchase of the print title. A library can lend such an item to its own patrons through regular circulation processes and to another library through interlibrary loan. Executing these functions online, therefore, creates no market harm beyond efficiencies created by faster loaning and return. At its most basic, format shifting is not transformative, nor is the circulation of a digital item. Both serve the same function and purpose as existed with the original print book. Under traditional 107 analysis, then, this project could fail. For society’s benefit, it should not.

Returning again to the Copyright Clause, the project above does nothing to harm incentive to authors, as authors still get paid for the creation of their works. There is no increase of the number of copies to the market, as the number of works circulated remains the same.

Courts should be encouraged to take a step back from Section 107 and look to the purpose of copyright in determining fair use. Failing that, they should be reminded that the four prongs in Section 107 are not exclusive. If not exclusive, then another factor should be introduced to the analysis: equity. Libraries’ resources are not unlimited and it makes no economic sense that they should have to spend scarce resources on the same material multiple times because of technology changes. They should be able to use the materials that they purchase fully in any technological era, regardless of the prevailing format.

⁶³ Id. at 226.

If both resetting the fair use test and the addition of a factor fails, libraries should then argue for a more nuanced test for the fourth factor. Instead of barring from fair use a copy that supplants or substitutes for the original for the same purpose, courts should adopt instead a test that sets a higher bar. For example, they could say that it is not fair use to create a copy that expands the number available in the market when the new copy which supplants or substitutes for the original for the same purpose. Or that mere substitution isn't enough but whether the substitution is unfair.

Under any of these three approaches, the project described above can flourish, for the benefit of society while still providing remuneration to authors. Books relevant to a researcher could be more easily identified and more quickly obtained. Books, whether out-of-print or within print, would be readily available to interested scholars, and law faculty and students nationwide would have access to the same resources, elevating the level of scholarship beyond. While there might be a market effect, due to technological efficiencies, it is not a market effect that thwarts the goals of copyright.

There likely would be resistance because of potential aftereffects. The first objection would be if libraries can do this, then why not the public at large? If the public at large can do this, then would the action invite a Napster-like community of pirates? Since all of the approaches above require control of the work and no additional copies to the market, this fear, while real, is fairly easily answered. Unless any actor controls the item digitized, they would be subject to a copyright infringement suit where the affirmative defense of fair use would fail.

The second is that if format shifting is permitted, would this disadvantage authors and provide a windfall to publishers who could then shift a book into a new format without negotiating new terms? Practically speaking, foreseeable uses have been a part of copyright law for some time, and since a series of cases in the late 1990s through early 2000s,⁶⁴ publishers and agents representing authors in negotiation have been particularly careful of licensing language restricting or permitting certain uses. As parties are always free to waive their fair use rights, should a publisher and author agree to limit publication to a given format, contract would prevail over any fair use claim.

This test brings copyright and fair use back full circle, to the Copyright Clause and its two principles to provide incentives to authors while stimulating societal development. The type of project described in this article arguably meets that test. It does not diminish the incentive for authors to create, as libraries will continue to select and buy materials. It just seeks to expand access to the materials in a manner that is consistent with current library services and practices.

Conclusion

⁶⁴ *Boosey & Hawkes Music Publishers, Ltd. v. Walt Disney Co.*, 145 F.3d 481 (1998) (noting that when “a license includes a grant of rights that is reasonably read to cover a new use (at least where the new use was foreseeable at the time of contracting), the burden of excluding the right to the new use will rest on the grantor”); *Random House, Inc. v. Rosetta Books, LLC*, 150 F.Supp.2d 613 (2001) (holding that restrictive terminology in a license agreement prevented new, foreseeable uses that fell outside of the accepted definition of that term)

While there are many factors to consider in library digitization projects --- costs, preservation, migration, integration with discovery platforms, document control, security, privacy --- copyright should not be one of the issues that prevents forward movement. In order to advance societal interests, libraries and universities should be willing to engage in activities designed to test fair use and challenge courts to recognize that even non-transformative, substitute uses can be fair.