

What Does the *HathiTrust* Decision Mean for Libraries?

Jonathan Band, policybandwidth

The library community welcomed the decision of the US Court of Appeals for the Second Circuit in *Authors Guild v. HathiTrust*, 755 F. 3d 87 (2d Cir. 2014). The decision has implications for libraries that go far beyond the specific facts of the case. This paper offers some preliminary thoughts on what these implications may be.

The broadest implication of the decision arises out of a footnote. Ever since the adoption of the library exceptions in 17 U.S.C. § 108, rights holders have argued that section 108 limits the availability of fair use to libraries, notwithstanding the savings clause in section 108(f)(4) that states explicitly that “nothing in this section in any way affects the right of fair use as provided by section 107.” In this litigation, the Authors Guild repeatedly argued that section 108 restricted fair use. Judge Baer rejected this argument in the district court, and the Second Circuit rejected it again in footnote 4. Citing the savings clause, the Second Circuit stated that “we do not construe § 108 as foreclosing our analysis of the Libraries’ activities under fair use....”¹ Thus, the decision holds unambiguously that libraries may take full advantage of the fair use right.

The *HathiTrust* decision also demonstrates how the fair use right applies in the context of a specific library activity: mass digitization. The decision clearly indicates that the acts of a library digitizing the works in its collection, and the library’s storage of the resulting digital files, are fair uses under section 107 of the Copyright Act. The decision, however, provides less certainty concerning the permissible access to those digital files. The only form of full-text access the decision addresses directly is access by the disabled. To be sure, this is an incredibly important result for these individuals. But the court provides little specific guidance concerning the permissibility of other forms of access. Nonetheless, the court’s more general pronouncements concerning fair use should be helpful to libraries trying to determine the range of permitted access to their mass-digitization projects.

1. Mass Digitization and Storage

The *HathiTrust* decision provides libraries with a clear roadmap for the digitization of the analog works in their collections. The court considered whether two uses of the digital copies deposited in the HathiTrust Digital Library (HDL) were fair use: full-text search and access for the print-disabled. The court found that “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.”² The court further found that HDL’s creating and maintaining four text-only copies of its entire database (one on the primary server at the University of Michigan, another at the mirror server at the University of Indiana, and two encrypted backup tapes at two secure locations on the University of Michigan campus) was justified for the purpose of balancing the load of user web traffic and as back-up in the case of a disaster.

Moreover, the court found that HDL demonstrated its need to retain image files in the four locations in addition to the text files. While the text files are required for text search and the creation of text-to-speech capabilities for the print-disabled, image files “provide an additional and often more useful method by which many disabled patrons...can obtain access to these works.”³

What is the practical effect of these holdings? Because providing full-text search capability justifies the creation and maintenance of a database of text files, a library could create and maintain a database of text files if the library provided full-text search capability of those text files. Likewise, because providing access for the print-disabled justifies the creation and maintenance of a database of image files, a library could create and maintain a database of image files if the library provided the print-disabled with access to those image files. Additionally, the library could create appropriate backup copies of these databases.⁴

As discussed below, the claims relating to “preservation” that were remanded to the district court actually concern the making of replacement copies “to be read and consumed by patrons.”⁵ The court distinguished the creation of replacement copies for patron consumption from the broader preservation function: “By storing digital copies of the books, the HDL preserves them for generations to come, and ensures that they will exist when the copyright terms lapse.”⁶ While the Second Circuit vacated the judgment concerning the “additional use” of making replacement copies, it left undisturbed the broader preservation function, thereby reinforcing the argument concerning the permissibility of mass digitization and storage.

In short, the *HathiTrust* decision indicates that a library could make digital copies of all the analog works in its collection, and store those copies as text and image files, if the library provided full-text-search capability and full-text access for the disabled.

2. Access to Works

While the court provided broad permission for libraries to digitize their collections, it addressed full-text access to the text and image files outside the library premises only in the context of the disabled. The court did not define the universe of the disabled entitled to full-text access, but the discussion of image files indicates that it goes well beyond the blind. The court stated that “[m]any legally blind patrons are capable of viewing these images if they are sufficiently magnified or if the color contrasts are increased.”⁷ The court then added that “other disabled patrons, whose physical impairments prevent them from turning pages or holding books, may also be able to use assistive devices to view all of the content contained in the image files for a book.”⁸

But the court provides little direct guidance with respect to access for readers **without** disabilities. Significantly, a search on HDL does not result in a display of snippets, in contrast to Google Book Search (GBS). Indeed, the court stressed that HDL “does not allow users to view any portion of the books they are searching.”⁹ Instead, HDL simply informs users which books contain the search term, and on what pages in those books the term appears. This three-judge panel of the Second Circuit evidently did not want to preclude the panel considering the pending appeal in the Google Books case from distinguishing the

HathiTrust decision on this basis. At the same time, this panel did not suggest that snippet display renders GBS unfair. This panel simply left the question open for the GBS panel. Accordingly, we will have to wait until the decision in the GBS case to know whether Judge Chin's holding permitting snippet display stands.

The court also did not resolve the issue of HDL's creation of a replacement copy "to be read and consumed by patrons" in the event a library's original copy is lost or destroyed. The court labeled this as a preservation use, and found that the plaintiffs may not have had standing to bring a claim related to this specific use. On this basis, the Second Circuit vacated the district court's judgment related to this use, and remanded the claim to the district court to determine whether any of the plaintiffs in fact had standing. On December 23, 2014, the parties signed a stipulation under which the plaintiffs dismissed this claim based on the defendant libraries' representation that they only made replacement copies from HDL where (i) the original copy that was being replaced was damaged, deteriorating, lost, or stolen; and (ii) it was determined, after reasonable effort, that an unused replacement could not be obtained at a fair price.

3. Suggestions Concerning Other Forms of Access

The *HathiTrust* decision could be read as suggesting that the display of thumbnails of digitized images (e.g., digitized archives of photographs) would be permissible. In support of the position that full-text search is a transformative use, the court cited two Ninth Circuit decisions that held that "the use of copyrighted thumbnail images in internet search results was transformative because thumbnail copies served a different function from the original copyrighted images."¹⁰

The court's more general fair use pronouncements could also assist libraries in determining what access to allow to their digitized copies. As the fair use jurisprudence has placed greater emphasis on whether a use is "transformative," rights holders have argued that the work itself had to be transformed in some manner for the use to be transformative. They pointed to parody as the quintessential transformative use.¹¹ In contrast, defendants have asserted that a use could be transformative if the use served a different purpose or function from the original, even if the work itself was not transformed. The copying of student papers for use by plagiarism-detection software is an example of this sort of use.¹²

In this case, the Authors Guild argued that the other circuits that had adopted this "functional transformation" approach had erred, that the Second Circuit itself had not adopted this approach,¹³ and that the Second Circuit should not follow the other circuits. Fortunately, the Second Circuit rejected the Authors Guild's assertion, finding that "a transformative work is one that serves a new and different function from the original work and is not a substitute for it."¹⁴ The court elaborated that "a use is transformative if it does something more than repackage or republish the original work."¹⁵ Because authors do not write with the purpose of enabling text searches of their books, the full-text search function does not "supersede the objectives or purposes of the original creation."¹⁶ The court added that "by enabling full-text search, the HDL adds to the original something new with a different purpose and a different character."¹⁷

In addition to clearly accepting the functional transformation approach, the court explained the impact of a finding of transformativeness on the fourth fair use factor, the effect of the use on the potential market: “under Factor Four, any economic ‘harm’ caused by transformative uses does not count because such uses, by definition, do not serve as substitutes for the original work.”¹⁸ The court later underscored this point: “The only market harms that count are the ones that are caused because the secondary use serves as a substitute for the original, not when the secondary use is transformative...”¹⁹ And again: “Lost licensing revenue counts under Factor Four only when the use serves as a substitute for the original and the full-text-search use does not.”²⁰

With respect to the third fair use factor, the amount and substantiality of the portion used, the court stressed that the extent of permissible copying varies with the purpose and character of the use. “For some purposes, it may be necessary to copy the entire copyrighted work, in which case Factor Three does not weigh against a finding of fair use.”²¹

In sum, the court found that a use is transformative if it has a different purpose or function from that of the original work. If the use is transformative, any economic harm such as lost licensing revenue is irrelevant for purposes of the fair use analysis. Further, the user may copy as much of the work as necessary to achieve his or her transformative purpose.

The court’s unambiguous endorsement of the functional transformation approach, combined with its discounting of lost revenue from such transformative uses, provides libraries with a solid basis for providing full-text access to its digitized copies when the underlying purpose of providing the access is different from the author’s original purpose. Thus, a library could provide access to digitized archival material such as an organization’s records. The purpose of the access is to enable scholarly research of the organization’s history, while the records were created to facilitate the operation of the organization. Although this reasoning probably would apply to most archival material, care should be exercised with respect to archives of entertainment content such as films. It is possible that many of the potential viewers of more popular films desire access to the films for entertainment rather than research purposes.

With respect to databases of digitized books, such as HDL itself, similar caution should be exercised. The Second Circuit found that providing the print-disabled with full-text access was not a transformative use: “By making copyright works available in formats accessible to the disabled, the HDL enables a larger audience to read those works, but the underlying purpose of the HDL’s use is the same as the author’s underlying purpose.”²² Even though the use was not transformative, the court nonetheless found that access was justified because of the special circumstances of the disabled. Because this rationale obviously would not apply if a library sought to provide full-text access for the non-disabled, the library typically would have to justify the access on functional transformation grounds. That is, the purpose of the use enabled by the access would have to be different from the author’s original purpose in writing the book.

For certain categories of books, identifying a different purpose should be straightforward. For example, books on the natural or physical sciences written before a certain date (say 1990) would be of interest to

readers for the purpose of researching the history of science, not for the purpose of advancing knowledge in that scientific field.

Librarians and scholars may be able to develop alternative approaches for identifying different purposes with respect to other categories of books. For example, under some circumstances consideration of library circulation statistics might be an appropriate basis for identifying a different purpose. If a novel circulated only five times in the 10 years prior to digitization, it could be safe to assume that anyone who sought access to the digitized copy of the novel did so for scholarly purposes such as studying the author or a literary movement. The circulation figures would act as a surrogate or proxy for the changing purpose of access to the work. The scholarly purpose could further be ensured if access were provided only to faculty and students, and not the general public.

Library circulation statistics could also provide a basis for determining whether to allow access to digitized books in certain situations the Second Circuit might consider non-transformative. (As noted above, the Second Circuit recognized that non-transformative uses could also be fair.) Many scholarly monographs on narrow topics might be borrowed a few times soon after publication, and then are hardly circulated at all. It is possible that a faculty member or graduate student who sought access to such a scholarly monograph after it was digitized would do so for the purpose intended by the author, i.e., to learn about the subject matter of the monograph. Nonetheless, because scholarship and research are among the legitimate purposes specifically enumerated in section 107, the first factor tilts in favor of the use, even though it is non-transformative. Moreover, the very low demand for the book, demonstrated by its infrequent circulation, suggests that there no longer is an actual or potential market for the book that could be harmed by providing digital access to the book. Accordingly, libraries perhaps should consider providing access to digitized books that have historically low circulation rates. If demand for a book (either digitally or in hard copy) increases after such access is provided, the library could decide whether it is appropriate to continue to allow full-text access to this re-discovered work.

4. Associational Standing

A final implication of the decision merits mention: the reduction of infringement actions by associations. The Second Circuit affirmed the district court's holding that associations such as the Authors Guild do not have standing to bring copyright infringement actions on behalf of their members. This holding will prevent US rights-holder associations from bringing broad, policy-driven copyright infringement actions. However, foreign associations that under foreign law have the right to bring enforcement actions on behalf of their foreign members would still have standing to bring copyright infringement actions in the United States on behalf of their members. Moreover, copyright class actions might still be permitted.

Conclusion

The *HathiTrust* decision provides libraries with much greater certainty concerning their mass-digitization projects. The decision indicates that the act of digitization, and the storage of digital files, is a fair use

if the libraries provide full-text search functionality and full-text access for disabled individuals. With respect to full-text access for the non-disabled outside of the library premises, the court ruled that a use could be transformative if the function or purpose of the use is different from that of the original work. This holding could reasonably be interpreted as permitting full-text access to most digitized archival material (except some popular entertainment materials). Full-text access to certain categories of digitized books, e.g., older books in certain scientific fields or books that are infrequently circulated, may also be permitted. Further discussion between librarians and scholars may lead to other approaches that could provide a basis for allowing full-text access to additional books.

Endnotes

- 1 775 F.3d at 94, n. 4.
- 2 *Id.* at 98.
- 3 *Id.* at 102–03.
- 4 The Second Circuit did not address the temporary copies made within the HathiTrust servers during the course of performing a search. Presumably the court considered these temporary copies to be so temporary as to not be copies within the meaning of the Copyright Act, see *Cartoon Network v. CSC Holdings*, 536 F.3d 121 (2nd Cir. 2008); or fair uses, see *Perfect 10 v. Amazon.com*, 508 F.3d 1146 (9th Cir. 2007). It is safe to assume that the court would also view any temporary copies made during the course of other “non-consumptive” uses of the database, such as text mining, as non-infringing.
- 5 *Id.* at 103.
- 6 *Id.*
- 7 *Id.*
- 8 *Id.*
- 9 *Id.* at 97.
- 10 *Id.* at 97.
- 11 See *Campbell v. Acuff-Rose*, 510 U.S. 569 (1994).
- 12 See *A.V. ex rel. Vanderhye v. iParadigms*, 562 F.3d 630 (4th Cir. 2009).
- 13 The Second Circuit had, in fact, already recognized functional transformation in *Bill Graham Archives v. Dorling Kindersley*, 448 F.2d 605 (2d Cir. 2006).
- 14 *HathiTrust* at 96.
- 15 *Id.*
- 16 *Id.* at 97.

17 *Id.*

18 *Id.* at 99.

19 *Id.*

20 *Id.* at 100.

21 *Id.* at 98.

22 *Id.* at 101.

© 2015 Jonathan Band



This article is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

To cite this article: Jonathan Band. “What Does the *HathiTrust* Decision Mean for Libraries?” *Research Library Issues: A Report from ARL, CNI, and SPARC*, no. 285 (2015): 7–13. <http://publications.arl.org/rli285/>.