

Alert Intellectual Property/Media

Second Circuit Holds Library Book-Scanning Project Is Fair Use

By Jonathan Bloom and
Sabrina Perelman

Introduction

On June 10, 2014, the Second Circuit took its first step into the intersection of online databases and copyright fair use when it held in *Authors Guild, Inc. v. HathiTrust*¹ that the creation of a full-text, searchable database of copyrighted books, and the provision of those books in formats accessible to individuals with print disabilities, are protected as fair use under section 107 of the Copyright Act. In affirming a ruling by the late Judge Harold Baer, Jr., the Court's ruling as to the searchable database relied primarily on what it found to be the highly transformative use made of the copyrighted works as well as on the fact that the word-search tool did not entail circulation of new copies of the books and thus did not substitute for the originals. In so ruling, the Second Circuit largely agreed with the district court's fair use analysis, but it clarified that simply adding value to a work (such as by converting it to a print-disabled accessible format) does not, by itself, render a use transformative, although in the case of print-disabled access, it found the lack of transformative value was not dispositive. The court also held that fair use also protected several complete copies made and retained to ensure the proper functioning of the database.

HathiTrust clearly bodes well for Google in the appeal involving its Google Books project, in which Google makes "snippets" of the text of books available as part of a searchable database of books in university library collections.² The decision accords with earlier search-engine fair-use rulings that have rejecting infringement claims arising out of knowledge-disseminating new digital uses of copyrighted works that do not intrude on the copyright owner's market.

Background

Beginning in 2004, several research universities, including the University of Michigan, the University of California at Berkeley, Cornell University, and the University of Indiana, contracted with Google to permit the search-engine giant to scan and create digital copies of the books in their collections. In October 2008, thirteen universities announced their plan to create a repository for these digital copies and organized HathiTrust to create and operate the HathiTrust Digital Library ("HDL"). HathiTrust currently has 80 member

institutions, and the HDL contains digital copies of more than 10 million works. HathiTrust permits three uses of the copyrighted works in the HDL:

1. *Searchable Database*: Allows the public to search across all digital copies in the repository for particular terms. Unless broader use is authorized, the search results show only the page numbers on which the search term is found and the number of times the term appears on each page.
2. *Access for Print-Disabled*: Allows member libraries to provide patrons with certified print disabilities access to the full text of the digitized works.
3. *Preservation*: Permits members to create a replacement copy of the work if the member already owns an original copy; the member's original copy was lost, destroyed, or stolen; and a replacement copy is not obtainable at a fair price elsewhere.

To facilitate these various uses, the HDL stores several digital copies of the works: one on its primary server in Michigan; another on its secondary server in Indiana; and two on separate backup tapes at the University of Michigan. Each copy contains the full text of the work, in a machine-readable format, as well as the images of each page in the work as they appear in the print version.

Separate from the HDL, the University of Michigan developed a project known as the Orphan Works Project (“OWP”),³ which was envisioned as a project to attempt to identify orphan works. If no copyright holder could be identified, the work would be made accessible in digital format to the OWP’s library patrons. However, before the OWP was ever brought online, concern that the screening process was not properly identifying orphan works led the university to suspend the project indefinitely.

The District Court Proceedings

The litigation was initiated in 2011 when twenty authors and authors’ associations sued HathiTrust, one of its member universities, and the presidents of four other member universities for copyright infringement in the Southern District of New York seeking declaratory and injunctive relief.⁴

The National Federation for the Blind and three print-disabled students were granted leave to intervene to defend their ability to continue using the HDL. The district court granted the defendants’ and intervenors’ motions for summary judgment on fair-use grounds.⁵

As a threshold matter, the court held that certain of the associational plaintiffs – the Authors Guild, the Australian Society of Authors Limited, and The Writers’ Union of Canada – lacked standing.⁶ These parties sued in their associational capacity on behalf of their members, which the district court held did not meet the statutory test for copyright standing.⁷

On the merits of the asserted fair use defenses, the court noted concerning the first fair-use factor (purpose and character of the use) that the HDL was undertaken to permit scholars to identify relevant works far more efficiently; to help university libraries to preserve their collections; and to provide print-disabled individuals access to library collections.⁸ The court found each of these uses to be transformative. With respect to the text-search database, the court found that the copies served “an entirely different purpose than the original works,” namely, “superior search capabilities rather than actual access to copyrighted material.”⁹ The court also found the print-disabled function transformative because it “was not the intended use of the original work”: “enjoyment and use by sighted persons.”¹⁰ Accordingly, the court found that the first factor tilted in the defendants’ favor on both challenged aspects of the HDL.¹¹

In light of the transformative nature of the uses, the court found that the second fair use factor (nature of the copyrighted work) was not dispositive, even though some 76 percent of the works in the HDL were fiction.¹² And because making entire copies of the works was necessary to fulfill the transformative purpose, the third factor – the amount and substantiality of the portion used – also favored the defendants.¹³

Finally, with respect to factor four – the impact on the market for or value of the copyrighted works – the court found that the plaintiffs had not shown a meaningful likelihood of future harm.¹⁴ It rejected the argument that each digital copy represented a lost sale because the purchase of an additional

copy “would not have allowed either full-text searches or access for the print-disabled individuals” – transformative uses that are “central” to the digitization project.¹⁵ The court also was unpersuaded by the plaintiffs’ concern with piracy, which it found to be unsupported by the record.¹⁶ Finally, the court found the purported undermining of plaintiffs’ “existing and emerging licensing opportunities” to be both conjecture and not cognizable because it related to a transformative market.¹⁷

On balance, the district court found that the HDL fell “safely within the protection of fair use,” noting that it “[could not] imagine a definition of fair use that would not encompass the transformative uses made by [the HDL] and would require that [the court] terminate this invaluable contribution to the progress of science and cultivation of the arts that at the same time effectuates the ideals espoused by the [Americans with Disabilities Act].”¹⁸

The court also held that the plaintiffs’ claims arising out of the OWP were not ripe.¹⁹

The Second Circuit’s Affirmance

The Second Circuit affirmed in a unanimous decision written by Judge Barrington D. Parker, although its fair-use analysis differed somewhat from that of the district court.

At the outset, the court of appeals agreed that the Authors Guild, Inc., the Australian Society of Authors Limited, and the Writers’ Union of Canada lacked standing because section 501 of the Copyright Act “does not permit copyright holders to choose third parties to bring suits on their behalf,”²⁰ whereas the other four authors’ associations had standing because foreign law conferred on them the exclusive right to enforce the copyrights of their foreign members.²¹

Turning to the merits, the court found both the full-text search database and access to the print-disabled to be fair use. With respect to the database, the court noted in evaluating the first factor that a transformative use is one that “does something more than repackage or republish the original copyrighted work.”²² Taking issue with an implication of the district court’s analysis, the court clarified that a use “does not become transformative by making an ‘invaluable

contribution to the progress of science and cultivation of the arts.”²³ The court explained that “[a]dded value or utility is not the test”; rather, a transformative work is one that “serves a new and different function from the original work and is not a substitute for it.”²⁴

Applying this standard, the court found the full-text search database to be “quintessentially transformative” because the result of a user’s word search is “different in purpose, character expression, meaning, and message from the page (and the book) from which it is drawn.”²⁵ As there was “no evidence that the [a]uthors write with the purpose of enabling text searches of their books,” and as the search results did not allow the users to view any portion of the books being searched, the court held that the “full-text search function does not ‘supersede the objects or purposes of the original creation.’”²⁶ Rather, by enabling the search function, the HDL “adds to the original something new with a different purpose and a different character.”²⁷

Indeed, the court noted that the full-text search function of the HDL added more to the copyrighted works than did the uses it found transformative in other cases, including *Cariou v. Prince* (also written by Judge Parker),²⁸ which involved Richard Prince’s collages incorporating plaintiff Patrick Cariou’s ethnographic photographs, and *Bill Graham Archives v. Dorling Kindersley Ltd.*,²⁹ which involved reduced-sized reproductions of concert posters in a biography of The Grateful Dead.³⁰ The court also cited as support cases involving the use of thumbnail images in Internet search results, which the Ninth Circuit held was transformative and fair use in *Perfect 10 Inc. v. Amazon.com, Inc.*³¹ and other cases.³² The court also relied on *A.V. ex rel. Vanderhye v. iParadigms, LLC*,³³ in which the Fourth Circuit held that the electronic copies of student papers made in connection with a plagiarism-detection computer program was transformative.³⁴ As in these cases, the court found that the different use to which the HDL put the plaintiffs’ works made it a transformative use.

As for the second fair use factor, the court stated that it “may be of limited usefulness” where creative works are being used for a transformative purpose and that it thus was not dispositive.³⁵

The court also found that the third factor weighed in favor of fair use, noting that for some purposes “it may be necessary to copy the entire copyrighted work”³⁶ and that the HDL full-text search database required the creation of complete digital copies to enable the full-text search to function.³⁷ The court rejected the plaintiffs’ argument that the creation and maintenance of copies in four different locations was excessive, finding that each was reasonably necessary to facilitate the HDL’s legitimate purposes: the copies on the “primary” server in Michigan and the “mirror” server in Indiana provided access to patrons while managing web traffic, and the additional two copies were encrypted backup tapes, disconnected from the Internet and placed in secure locations so that data could be restored in the event of large-scale data loss to the two servers.³⁸

Turning to market harm, which the court called the “most important element of fair use,”³⁹ the court emphasized that it concerns harm arising from the secondary use serving as a substitute for the original work.⁴⁰ Accordingly, any economic harm caused by a transformative use “does not count because such uses, by definition, do not serve as substitutes for the original work.”⁴¹ The court proceeded to reject both of the plaintiffs’ market-harm arguments. As for the argument that the HDL impaired the development of a market for licensing books for digital search, the court held that because the search function did not serve as a substitute for the books being searched, the loss of this potential licensing market was not cognizable market harm.⁴² And as for the argument that the HDL created the risk of a security breach that might lead to widespread unauthorized copying, the court cited record evidence that the defendants had undertaken extensive security measures and concluded that the posited potential harm was speculative.⁴³

In sum, the court held that the HDL’s digitizing of copyrighted works for the purpose of permitting full-text searches was fair use.

Addressing print-disabled access, the Second Circuit acknowledged that by making copyrighted works available in formats accessible to the disabled, the HDL enables a larger audience to read the works.⁴⁴ But the court emphasized that this was not

transformative because the underlying purpose of the HDL’s use was the same as the author’s original purpose.⁴⁵ The court analogized the conversion of works into a print-disabled-accessible format to the translation of the works for non-English-speaking audiences. In neither case is the resulting converted work transformative “simply because it enables a new audience to read the work.”⁴⁶

The court noted, though, that transformative value “is not absolutely necessary for a finding of fair use.”⁴⁷ (As we discussed in a prior Alert, the Second Circuit made a similar finding in *Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.*, – F.3d –, 2014 WL 2219162 (2d Cir. May 30, 2014), with respect to an unauthorized tape recording of a newsworthy earnings call.) The court pointed out that the legislative history of the 1976 Copyright Act identifies making copies accessible for the blind as a “special instance illustrating the application of the fair use doctrine”⁴⁸ and concluded that the “unique circumstances presented by print-disabled readers” warranted a finding that the first factor favored fair use.⁴⁹

The court found that factor two tipped against fair use but that this did not preclude a finding of fair use.⁵⁰

With respect to the third factor, the plaintiffs claimed that HDL’s retention of digital image files as well as text-only files was excessive copying, but the Second Circuit disagreed, holding that the text files were “required for text searching and to create text-to-speech capabilities” and that the image files would “provide an additional and often more useful method” by which disabled patrons, especially students and scholars, could access the works, including pictures, charts, diagrams, and layout, which are not captured in a text-only copy.⁵¹

As for market harm, the court noted that because the market for books accessible to the handicapped was so insignificant, it is common for authors to forgo royalties from such books and that, accordingly, this factor favored the defendants.⁵²

Thus, notwithstanding its finding that the use was not transformative, the court held that the HDL’s print-disabled access function was fair use.⁵³

Discussion

HathiTrust is an initial indication of the Second Circuit's approach to fair use as applied to a text-searchable database. The court's endorsement of the HDL is consistent with the fair-use rulings of other appellate courts in cases involving search engines and other digital informational tools. Notably, though, the court emphasized that "[a]dded value or utility is not the test,"⁵⁴ rejecting the notion that merely adding functionality that facilitates access to or use of a work is, by itself, transformative. Instead, the court stressed that a transformative work is one that "serves a new and different function from the original work and is not a substitute for it,"⁵⁵ and it based its determination that the HDL search function was transformative on (i) there being "little or no resemblance between the original text and the results of the HDL full-text search"⁵⁶ and (ii) the fact that the full-text search function "does not 'supersede[] the objects [or purposes] of the original creation."⁵⁷ By ruling, on the other hand, that merely increasing access for the print-disabled was *not* transformative, the court rejected a theory that could have threatened the ability of copyright owners to control digital dissemination of their works online.

The Second Circuit's focus on the new and non-superseding purpose of the HDL database – and hence its transformativeness – was echoed in its emphasis on usurpation of the copyright owner's market being the gravamen of cognizable market harm and the corollary that copyright owners cannot control transformative uses, which would appear to present a difficult hurdle for the Google Books plaintiffs.

HathiTrust is surely an indication of how the Second Circuit is likely to rule in the Google Books case (*Authors Guild v. Google*), now being briefed on appeal.⁵⁸ As noted, that case concerns Google's creation of digital copies of books in the collections of participating libraries, which Google catalogs and makes available for a full-text search.⁵⁹ One difference between Google Books and HDL is that Google Books search results return a "snippet" of the book itself, while the HDL provides only a list of the pages on which the search term appears.⁶⁰ In the district court's summary judgment ruling in favor of Google,

the court (per Second Circuit Judge Denny Chin, who retained the case after being elevated to the court of appeals) emphasized the transformative nature of Google Books, finding that it:

- "transforms expressive text into a comprehensive word index that helps readers, scholars, researchers, and others find books";⁶¹
- "has become an important tool for libraries and librarians and cite-checkers as it helps to identify and find books";⁶² and
- "has transformed book text into data for purposes of substantive research, including data mining and text mining in new areas, thereby opening up new fields of research."⁶³

The court concluded that Google Books "preserves books, in particular out-of-print and old books that have been forgotten in the bowels of libraries, and it gives them new life. It facilitates access to books for print-disabled and remote or underserved populations. It generates new audiences and creates new sources of income for authors and publishers. Indeed, all society benefits."⁶⁴

It will be interesting to see how the Second Circuit evaluates Judge Chin's focus on the functional use of the text as a predicate for finding transformative value. The court of appeals' finding in *HathiTrust* that the HDL's full-text search function is a new and transformative use suggests it will be receptive to a theory of functional transformativeness based on the repurposing of the original text. But perhaps because the HDL search-results page does not incorporate any of the text of the copyrighted works (unlike the text "snippets" available on Google Books), the *HathiTrust* court did not rely on the use of copyrighted text in a functional manner as an information-location tool, instead relying on the different appearance of the HDL search results from the original text and the different, non-superseding purpose of the HDL search function.

While transformativeness occupies a central position in the fair-use inquiry, as the *HathiTrust* court pointed out in discussing print-disabled access, fair use does not always require a finding of transformative value. In *Swatch*, similarly, the Second Circuit relied in its factor-one analysis on the "important public purpose"

of defendant Bloomberg's copying of an earnings call for news reporting, which, the court held, outweighed the lack of transformative value.⁶⁵ (The *Swatch* court also relied on the lack of an actual or likely potential market for the earnings call.) These rulings indicate that when a nontransformative use nevertheless advances an important public interest without supplanting the copyright owner's market, wholesale copying does not necessarily preclude fair use.

Conclusion

The Second Circuit's interpretation of fair use as applied to a searchable database bodes well for Google, as it suggests that the Second Circuit will look favorably on digitization that serves a different purpose from the original copyrighted work and does not supplant either the existing or a likely future market for the originals.

1. *Authors Guild, Inc., et al. v. HathiTrust, et al.*, – F.3d –, No. 12-4547-cv, 2014 WL 2576342 (2d Cir. June 10, 2014) (“*HathiTrust*”).
2. See *Authors Guild v. Google, Inc.*, 13-4829-cv (2d Cir.), on appeal from *Authors Guild, Inc. v. Google, Inc.*, 954 F. Supp. 2d 282 (S.D.N.Y. 2013) (Chin, J.).
3. An “orphan work” is an out-of-print work that is still in copyright but whose copyright holder cannot be readily identified or located. *HathiTrust*, 2014 WL 2576342, at *2.
4. See *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 (S.D.N.Y.).
5. *Authors Guild, Inc. v. HathiTrust*, 902 F. Supp. 2d 445, 458-64 (S.D.N.Y. 2012).
6. *Id.* at 453-54.
7. *Id.*
8. *Id.* at 459.
9. *Id.* at 460.
10. *Id.* at 461.
11. *Id.* at 459-61.
12. *Id.* at 461-62.
13. *Id.* at 462.
14. *Id.* at 462-64 (“Where a use is noncommercial, as it is here, the plaintiff must show by a preponderance of the evidence that some meaningful likelihood of future harm exists, a test Plaintiffs fail at least on this fact pattern.”) (internal quotation marks and citations omitted).
15. *Id.* at 462.
16. *Id.*
17. *Id.* Although the plaintiffs had asserted that the Copyright Clearance Center might have eventually developed a license for the uses to which the defendants put the works, the court found that the CCC had no plans to provide for or develop such a license.” *Id.* And while the plaintiffs contended that Congress might eventually “find a way to regulate this area of the law,” the court noted that “it is not the court’s job to apply laws that have not yet been written.” *Id.* (quoting *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 456 (1984)).
18. *Id.* at 464. The district court also found that the HDL was protected under the Chafee Amendment, which permits an “authorized entity” to reproduce or distribute copies of a previously published, non-dramatic literary work in specialized formats exclusively for the use by the blind or other persons with disabilities. *Id.* at 465; see also 17 U.S.C. § 121.
19. *Id.* at 455-56.
20. *HathiTrust*, 2014 WL 2576342, at *4 (quoting *ABKCO Music, Inc. v. Harrisongs Music, Ltd.*, 944 F.2d 971, 980 (2d Cir. 1991)).
21. *Id.* (citing *Itar-Tass Russian News Agency v. Russian Kurier, Inc.*, 153 F.3d 82, 93-94 (2d Cir. 1998)).
22. *Id.* at *6.
23. *Id.* (quoting *HathiTrust*, 902 F. Supp. 2d at 464).
24. *Id.*
25. *Id.* at *7.
26. *Id.* (quoting *Campbell*, 510 U.S. at 579) (alterations omitted).
27. *Id.*
28. 714 F.3d 694 (2d Cir. 2013).
29. See 448 F.3d 605 (2d Cir. 2006).
30. *HathiTrust*, 2014 WL 2576342, at *7.
31. 508 F.3d 1146, 1165 (9th Cir. 2007).
32. *HathiTrust*, 2014 WL 2576342, at *8.
33. 562 F.3d 630, 639-40 (4th Cir. 2009).
34. See *HathiTrust*, 2014 WL 2576342, at *8.
35. *Id.* (quoting *Cariou v. Prince*, 714 F.3d 694, 706 (2d Cir. 2013)).
36. *Id.*
37. *Id.*
38. *Id.* at *8-9.

39. *Id.* at *7 (citations omitted).
40. *Id.* at *9.
41. *Id.* (citing *Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 15 605, 614 (2d Cir. 2006)).
42. *Id.* at *9-10.
43. *See id.* at *10-11.
44. *Id.* at *11.
45. *Id.* at *11-12.
46. *Id.*
47. *Id.* (quoting *Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.*, – F.3d –, 2014 WL 2219162, at *7 (2d Cir. 2014)).
48. *Id.* (quoting H.R. Rep. No. 94-1476, at 73 (1976), *reprinted in* 1976 U.S.C.C.A.N. 5659, 5686).
49. *Id.*
50. *Id.*
51. *Id.* at *13.
52. *Id.* at 30 (quotation omitted). In light of the court's finding that the two uses in question constituted fair use under section 107 of the Copyright Act, it found it unnecessary to consider whether the disability-access use also is protected under the Chafee Amendment. *Id.* at *13 n.7.
53. The court likewise affirmed the district court's holding that the plaintiffs' claims concerning the orphan work project were not ripe. The court noted that it was "far from clear that" the University of Michigan or HathiTrust would reinstate the OWP in a manner that would infringe the copyrights of any proper plaintiffs and that it thus could not be said that any of the plaintiffs faced a "certainly impending harm." *Id.* at *15.
54. *Id.* at *6.
55. *Id.*
56. *Id.* at *7.
57. *Id.*
58. The authors' appellate brief was filed on April 7, 2014, and Google's appellee brief is due to be filed on or before July 3, 2014.
59. *See Google Books*, 954 F. Supp. 2d at 285-87.
60. *Id.* at 291.
61. *Id.* at 291.
62. *Id.*
63. *Id.*
64. *Id.* at 293.
65. *Swatch Grp. Mgmt. Servs. Ltd.*, 2014 WL 2219162, at *16.

If you have questions concerning the contents of this issue, or would like more information about Weil's IP/Media practice group, please speak to your regular contact at Weil, or to the editors or practice group members listed below:

Editors:

Randi Singer (NY)	Bio Page	randi.singer@weil.com	+1 212 310 8152
Jonathan Bloom (NY)	Bio Page	jonathan.bloom@weil.com	+1 212 310 8775

Contributing Authors:

Jonathan Bloom (NY)	Bio Page	jonathan.bloom@weil.com	+1 212 310 8775
Sabrina Perelman (NY)	Bio Page	sabrina.perelman@weil.com	+1 212 310 8752

© 2014 Weil, Gotshal & Manges LLP. All rights reserved. Quotation with attribution is permitted. This publication provides general information and should not be used or taken as legal advice for specific situations that depend on the evaluation of precise factual circumstances. The views expressed in these articles reflect those of the authors and not necessarily the views of Weil, Gotshal & Manges LLP. If you would like to add a colleague to our mailing list, please [click here](#). If you need to change or remove your name from our mailing list, send an email to weil.alerts@weil.com.