FAIR USE & MASS DIGITIZATION: THE FUTURE OF COPY-DEPENDENT TECHNOLOGIES AFTER *AUTHORS GUILD V. HATHITRUST*

Angel Siegfried Diaz†

From the wisdom contained in oral tradition to the artistic breakthroughs stored and shared on the cloud, societies continually benefit from preservation and access to the knowledge contained in original content. The Digital Age has enabled individuals around the world to store, organize, and share everything from photographs and videos to their academic writing and correspondence. Yet, for all of the information accessible via the Internet, the archived collections of university libraries remain out of reach to many, and the price of digitization remains largely cost-prohibitive.[1] In 2004, Google Inc. announced its intention to scan and digitize the collections of several leading university libraries.[2] The project would create a product that allows users to search across the entire corpus of scanned works.[3] For works still under copyright, the results page would present the user with the word that was queried in addition to some of the phrases surrounding it in order to give the user context.[4] The results page would also provide links to where the work could be legally purchased. Rights holders that did not wish to participate in the project could opt-out from the project, but works in the public domain would be fully viewable and available for download.[5] The universities partnering with Google and the Internet Archive established HathiTrust, an organization that would facilitate "collective action on a grand

† J.D. Candidate, 2014, University of California, Berkeley School of Law.

1. *See* Peter Menell, Knowledge Accessibility and Preservation Policy for the Digital Age, 44 HOUS. L. REV. 1013, 1015 n.2 (2007) ("The University of Michigan estimates that, using the technology and resources at its disposal, it would have taken more than 1,000 years to fully digitize its seven million volume collection."). Some European countries have undergone similar efforts with success, but none have undergone a project at the scale of the Google Books project.

2. Press Release, GOOGLE, Google Checks out Library Books (Dec. 14, 2004), http://googlepress.blogspot.com/2004/12/google-checks-out-library-books.html.

3. GOOGLE BOOKS LIBRARY PROJECT, books.google.com/googlebooks/library/index.html (last visited Mar. 28, 2013).

4. *Id.*

5. *Id.*

scale" and engage in projects that put the digital scans from Google to novel uses.[6]

However, several rights holders complained that the unauthorized scanning of their works infringed their copyright and filed suit against Google and HathiTrust. Now, after eight years of waiting and speculation, the case against HathiTrust resulted in the first opinion to apply a fair use analysis to unauthorized scanning for purposes of creating a full-text search index.[7] Largely aligning itself with the increasing body of case law privileging technologies that serve educational and research purposes, the court held that HathiTrust's index was a transformative use that did not serve the same expressive purposes as the underlying copyrighted works.[8] Additionally, the court held that HathiTrust's use of the digital scans to make copies for the print-disabled[9] was fair; it served a market not contemplated by the copyright holder and was also privileged under § 121 of the Copyright Act.[10]

This Note discusses the future of digital libraries and other products reliant on mass digitization in the wake of the *HathiTrust* decision. First, this Note presents an overview of U.S. copyright protection and the ways in which its goal of incentivizing authors has consistently been balanced by efforts to protect preservation, access, and fair use. Next, the Note explores early formalities and the establishment of the national archive before discussing the modern Copyright Act and its series of statutory exemptions. After outlining the statutory provisions, the Note then discusses the fair use doctrine, focusing on its application to cases involving use for educational and research purposes.

Second, this Note discusses the trial court opinion in *Authors Guild, Inc. v. HathiTrust* and the court's fair use finding regarding the full-text search index and copies for the print disabled. The Note argues that this decision rightly solidifies the growing body of jurisprudence protecting uses that have a transformative purpose and that pose no threat of artistic substitution while providing a public benefit. However, the Note also argues that the same consideration cannot be applied to HathiTrust's copies for the print-disabled

6. *See HathiTrust's Past, Present, and Future*, HATHITRUST DIGITAL LIBRARY, http://www.hathitrust.org/blogs/perspectives-from-hathitrust/hathitrusts-past-present-and-future (last visited Mar. 28, 2013).

7. *See* Authors Guild, Inc. v. HathiTrust, No. 11 CV-4351 (HB), 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012).

8. *Id.* at *14.

9. This Note uses "print-disabled" to refer to individuals that cannot read printed materials due to visual or physical disabilities.

10. *HathiTrust*, 2012 WL 4808939, at *15.

and that the trial court's fair use application and interpretation of § 121 may be narrowed or corrected on appeal.[11]

Third, this Note discusses the *HathiTrust* decision's effect on the future of the Google Books case and argues that the fair use ruling paves the road for a similar finding while also giving Google leverage in its ongoing settlement negotiations.[12] The Note argues that Google's fair use defense is equally as strong as HathiTrust's despite Google's commercial exploitation because no artistic substitution exists and Google Books helps rather than harms the market for the copyrighted works.

Fourth, after exploring the judicial efforts to protect useful technologies as a matter of public policy, this Note explores legislative solutions that would better advance copyright's goals of promoting education, research, preservation, and access. The Note advocates for (1) expansion of § 108 library privileges to cover mass digitization and noncommercial exploitation; (2) expansion of protected entities that may make copies for the print-disabled under § 121; (3) establishment of a regulatory board to govern the exploitation of orphan works; and (4) categorical protection for technological uses that make use of copyrighted works in a way that does not substitute for the original's expressive content.

I. AN OVERVIEW OF U.S. COPYRIGHT LAW: PROMOTING THE USEFUL ARTS AND SCIENCES VIA ACCESS, PRESERVATION, AND FAIR USE

The U.S. Constitution grants Congress the power to "promote the Progress of Science and useful Arts" by granting authors and inventors "exclusive [rights] to their respective Writings and Discoveries."[13] While the clause focuses on incentivizing authors, its purpose is much broader. The founders also intended to promote progress by facilitating learning through public access and a protected public domain.[14] This Section traces the development of U.S. copyright protection. From the formalities that premised early copyright laws to the statutory exemptions scattered across

11. A copy of the Plaintiff's February 25, 2013, opening appellate brief is available at http://thepublicindex.org/docs/cases/hathitrust-2ndcir/41-ag-brief.pdf.

12. *See* Compl., Author's Guild v. Google Inc., 770 F. Supp. 2d 666 (S.D.N.Y. 2011), *available at* https://www.eff.org/node/53647.

13. U.S. CONST. art. I §8.

14. *See* L. Ray Patterson & Craig Joyce, *Copyright in 1791: An Essay Concerning the Founders' View of the Copyright Power Granted to Congress in Article I, Section 8, Clause 8 of the U.S. Constitution*, 52 EMORY L.J. 909, 946 (2003).

the modern Copyright Act, access, preservation, and fair use are considerations that persistently inform the scope of protection and liability.

A. EARLY U.S. COPYRIGHT LAW: THE ROLE OF FORMALITIES AND CATALOGUING

1. The Deposit Requirement and the Smithsonian Institution Promote Preservation and Access to Copyrighted Work

Beginning with the earliest versions of the Copyright Act, statutory requirements sought to promote preservation and access to copyrighted works through cataloging and with the establishment of a Library of Congress. Under the 1790 Act, authors of registered works were required to publish a copy of their registration "in one or more newspapers" within two weeks of registration and were required to run the notice "for the space of four weeks."[15] Moreover, authors forfeited copyright protection unless they deposited a printed copy of their work "in the clerk's office of the district court where the author or proprietor shall reside."[16] Primarily intended to defeat future claims of authorship, the notice and deposit requirement also became a way to ensure copyrighted works were archived and accessible to the public.[17] Shortly thereafter, Congress established the Smithsonian Institution[18] and expanded the deposit requirement by providing that a copy of copyrighted works needed to be deposited with the newly formed Smithsonian as well as with the Librarian of Congress[19] within three months of publication.

Unlike the district court's deposit requirement, the Smithsonian and Librarian of Congress copies were not necessarily preconditions to copyright.[20] Consequently, the copies for the Smithsonian and Librarian of Congress were primarily aimed at expanding the national collection for preservation and public access. To that end, Congress made it easier for authors to deposit their works by allowing them to mail their copies free of postage.[21] Unfortunately, while new works began flooding into the Smithsonian, few of them came from major publishers of creative and scholarly works.[22] Instead, most of them were works such as textbooks and

15. Act of May 31, 1790, ch. 15, § 3, 1 Stat. 124, 125 (1790).
16. *Id.*
17. Menell, *supra* note 1, at 1026–27.
18. Act of Aug. 10, 1846, ch. 178, § 10, 9 Stat. 102, 106 (1851).
19. *Id.*
20. *See* Jollie v. Jacques, 13 F. Cas. 910, 912 (S.D.N.Y. 1850).
21. *See* Act of Mar. 3, 1855, ch. 201, § 5, 10 Stat. 683, 685.
22. *See* Menell, *supra* note 1, at 1027.

prints, which carried little archival value and were difficult to store.[23] A legislative overhaul of the national library aimed to tackle the early inefficiencies of the archival initiatives by repealing the deposit requirement and transferring copyright records and deposits to the Department of the Interior.[24] Authors were no longer required to deposit a copy of their works with the district court, only with the Librarian of Congress.

2. Transforming the Library of Congress into a National Repository

With later amendments to copyright law, Congress sought to better enable the Library of Congress to serve as a repository of the nation's works that would be available to the public. In 1865, Congress empowered the Librarian of Congress to make written demand of authors who failed to comply with the deposit requirement within one month of registration.[25] Authors that did not comply with the deposit requirement by one month after receiving the written demand lost copyright protection.[26] Five years later, Congress added a monetary fine for failure to submit a copy of registered works.[27] By 1870, Congress centralized authority for all copyright business with the Librarian of Congress[28] and required the office to create a catalogue of deposited works in addition to an annual report to Congress.[29] The 1870 Act specified that authors must deposit a copy of their work within ten days of registration to receive copyright protection.[30] The Act also continued to issue free postage for deposits[31] and imposed a fine for failure to comply with the deadline.[32] These changes allowed the Library of Congress to quickly swell in size as new works poured into the collection.

This Congressional update coincided with a shift in university education that stressed original research—a change that also increased the amount of new works just as the stricter deposit requirement was expanding the size of

23. *See* Menell, *supra* note 1 (citing Smithsonian Institution Annual Report 40 (1856)).

24. *See* Menell, *supra* note 1, at 1027 (citing Act of Feb. 5, 1859, ch. 22, §§ 5–6, 11 Stat. 379, 280.

25. Copyright Act of 1856, ch. 126, § 3, 13 Stat. 540, 540 *amended by* Copyright Act of 1909, ch. 329, § 63, 35 Stat. 1075, 1088 (1856).

26. *Id.*

27. *See* Act of Feb. 18, 1867, ch. 43, § 1, 14 Stat. 395, 395 (1867) (empowering the Library of Congress to impose a $25 fine on authors that fail to deposit a work within one month of publication).

28. *See* Act of July 8, 1870, ch. 230, §85, 16 Stat. 212.

29. CONG. GLOBE, 41st Cong., 2d Sess. 2684 (1870).

30. *Id.* at 2683.

31. Act of July 8, 1870, ch. 230, § 95, 16 Stat. 212.

32. *Id.* § 94.

the national collection.[33] Predicting that the educational shift would increase demand for access to the national collection, Congress amended the Copyright Act in 1891 and required the Library of Congress to publish catalogues of their collections at regular intervals to enable the public to conduct research.[34] Despite their limited utility, the late 19th century amendments to U.S. copyright law reflected Congress's growing focus on public benefit instead of private monopoly.

B. MODERN U.S. COPYRIGHT LAW: STATUTORY EXEMPTIONS AND THE FAIR USE SAFETY VALVE

Copyright law has changed dramatically since the turn of the 20th century. In 1909, Congress made the first significant changes to the copyright law by enacting a new copyright act ("1909 Copyright Act"). The 1909 Copyright Act was the 20th century's first major revamp of copyright law. In the legislative history leading up to the Act, the House Committee on Patents Report noted that Congress designed copyright law "[n]ot primarily for the benefit of the author, but primarily for the benefit of the public."[35] While the 1909 Copyright Act was the first attempt to modernize U.S. copyright law, its effect was less than dramatic. Moreover, the proliferation of new technologies and the increasing international market for copyrighted works created new concerns and questions about the scope of copyright protection. These new questions regarding the scope of copyright protection led the Copyright Office to fund a series of thirty-five studies to investigate the various functions of U.S. copyright law and its differences from the international community.[36] However, from the growing influence of motion picture and television industry to the possibilities and challenges presented by photocopying technology, competing interests stalled Congressional reaction to the Copyright Office's various studies.[37]

The studies and negotiations eventually produced the 1976 Copyright Act. This Section explores the provisions of the 1976 Copyright Act, subsequent updates to the Act, and the role of the fair use doctrine. The 1976 Copyright Act largely maintained the features of the 1909 Copyright Act but reinforced the importance of preservation and access for research

33. *See* Menell, *supra* note 1, at 1028–29.

34. *Id.* at 1029.

35. H.R. Rep. No. 60-2222, at 7 (1909).

36. *See* Harry G. Henn & Walter J. Derenberg, *Introduction* to 1 STUDIES ON COPYRIGHT ix, ix (Arthur Fisher Memorial ed. 1963).

37. *See* Menell, *supra* note 1 (quoting H.R. Rep. No. 89-2237, at 31–32 (1966) (discussing the inability of current copyright law to accommodate new technologies)).

purposes by maintaining the deposit requirements,[38] creating the Television and Radio Archives within the Library of Congress,[39] and establishing statutory protections for libraries that allowed them to engage in non-commercial photocopying.[40] Later, the spread of digital technology led Congress to once again create a series of statutory protections for preservation and access under the DMCA.[41] Lastly, the fair use doctrine—first codified in the 1976 Copyright Act—continues to function as a safety valve for uses of copyrighted works that promote education, research, and archival purposes.

1. The Modern Role of the Copyright Office: A Focus on Preservation and Archiving

While the 1976 Copyright Act kept the deposit requirement, it no longer preconditioned copyright protection on deposit.[42] Instead, Congress focused on the requirement's successful effect on preservation and attempted to ensure the continuous flow of new works into the national archive by continuing to impose fines for failure to comply with the deposit requirement: $250 per work and an additional $2,500 for willful violations.[43] Similarly, the 1976 Act continued to charge the Copyright Office with maintaining and publishing a catalog of copyrighted works and instructed the Register of Copyrights to produce catalogs "on the basis of practicability and usefulness" for each individual medium.[44]

2. Section 108 Protections for Libraries & the Expansion of the National Archives

As the role of formalities shifted and the national archive faced problems maintaining an exhaustive and usable record, private and public universities began making use of new technological tools that help their patrons get better access to works in the libraries' collections. Congress encouraged this behavior by enacting §§ 108 and 121 of the 1976 Copyright Act, which

38. *See* Copyright Act of 1976, § 101, 90 Stat. 2579.

39. *See* American Television and Radio Archives, Pub. L. No. 94-553, § 113, 90 Stat. 2541, 2601 (1976).

40. *See* Copyright Act of 1976, § 108, 90 Stat. 2541.

41. *See* Digital Millennium Copyright Act, Pub. L. 105-304, § 404, 112 Stat. 2860, 2890 (1998).

42. *See* Copyright Act of 1976, § 101, 90 Stat. at 2579 (current version at 17 U.S.C. §407 (2000)).

43. *See id.* (current version at 17 U.S.C. § 407(d) (2000)).

44. Copyright Act of 1976, Pub. L. No. 94-553, § 101, 90 Stat. 2541, 2592–93 (current version at 17 U.S.C. § 707(a) (2000)).

created a series of statutory exemptions withholding copyright liability for libraries.

The § 108 exemption grows out of the "photocopying revolution," which held the promise of increasing access and knowledge but also presented the possibility that mass replication would harm the market for copyrighted works.[45] Rather than subject library photocopying to uneven application under the fair use doctrine, discussed *infra* Section I.B.3, Congress established a detailed statutory exemption that allows libraries to engage in non-commercial photocopying without fear of direct or indirect liability.[46] This statutory exemption represents Congress's effort to balance the social utility of technological advancement with the need to protect authors.

Libraries and archives seeking § 108 protection must follow the statute's itemized requirements.[47] First, the photocopying must be done "without any purpose of direct or indirect commercial advantage."[48] Second, the libraries and archives must be open to the public, or at least to persons doing research in the specialized field that is the subject of the collection.[49] This requirement incentivizes libraries of all kinds to become more accessible, promoting the Congressional goal of ensuring that copyright remains "for the benefit of the public."[50] Third, all copies and distributions of copyrighted works are required to contain a notice of copyright.[51] Moreover, while § 108(d)(1) presumes that the library requests are for research purposes, it requires libraries to turn away patrons suspected of a different purpose.[52]

On balance, § 108 and the expansion of the national archive reflect the potential utility of allowing libraries and archives to make use of developing technologies in ways that enrich their collections and make them more useful to the public without harming the market for the copyrighted works. This

45. *See* 2-8 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT, § 8.03 (1992); *see also* 17 U.S.C. § 108.

46. *Id.*

47. Copyright Act of 1976, §108, Pub. L. No.94-553, 90 Stat. at 2546.

48. *Id.* at § 108(a)(1). Notably, the meaning of "commercial intent" refers to the purpose of the reproduction or distribution, not to the nature of the library. This expands the scope of the exemption beyond non-profit libraries to also include those belonging to businesses and proprietary institutions. Within the language of the statute, "commercial purpose" refers to the motivation for the act of copying or distribution itself. Consequently, even if a proprietary institution's library makes photocopies that are used for a commercial service, the copying is protected.

49. *Id.* at § 108(a)(2).

50. *See* H.R. Rep. No. 60-2222, at 7 (1909).

51. Copyright Act of 1976, §108.

52. *Id.*

balance of competing interests continued to inform subsequent efforts to empower libraries to preserve and disseminate knowledge.[53]

In addition to § 108, the first Congressional attempt to regulate copyright through the DMCA similarly exempted libraries from the DMCA's anti-circumvention provision.[54] Congress intended this exemption, codified in § 1201, to protect copyright owners by prohibiting the evasion of technological steps taken to restrict access to digital copies.[55] However, § 1201 exempts library circumvention only to the extent the activity is done to help the library decide whether or not to include the work in its collection.[56]

3. Fair Use and Non-Expressive Use

In addition to the statutory exemptions carved out in the 1976 Copyright Act, Congress also codified the fair use doctrine, a long-used affirmative defense to copyright infringement.[57] The preamble to the statute further promotes the goals of access and education by directing courts to favor uses related to "criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research."[58] In its latest treatment of the doctrine, the Supreme Court held that fair use is intended to ensure copyright law does not "stifle the very creativity which the law is intended to foster."[59]

Section 107 requires courts to consider four factors when evaluating a fair use defense:

53. For example, in 1998, Congress added twenty years to the term of copyright protection in the Sonny Bono Copyright Term Extension Act. This longer period of protection was balanced by an extension of library privileges under§ 108(h). *See* The Sonny Bono Copyright Term Extension Act, Pub. L. 105-298 (1998). The exemption in §108(h) allowed libraries and archives to reproduce and distribute copies of out of print and orphan works during the last twenty years as long as the use was for purposes of research and preservation. But Congress sought to further protect copyright owners by only granting the exemption after libraries took reasonable steps to conclude that 1) the work was not being commercially exploited, 2) could not be secured at a fair price, and that 3) the copyright holder had not informed the Register of Copyright where the work was available.

54. Digital Millennium Copyright Act, Pub. L. 105-304, § 103, 112 Stat. at 2866 (1998) (current version available at 17 U.S.C. § 1201(d) (2006)).

55. *Id.*

56. *Id.*

57. Copyright Act of 1976, § 107, Pub. L. No.94-553, 90 Stat. at 2546 (current version available at 17 U.S.C. § 107 (2006)). Codification was not intended to "change, narrow, or enlarge [the doctrine] in any way."). 4-13 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05 (1992).

58. 17 U.S.C. § 107 (2006).

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

> (1) the purpose and character of the use, including whether the use is of commercial nature or is of a nonprofit educational purpose;
>
> (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.[60]

These four fair use factors are meant to be illustrative guideposts for trial courts, but they are not exhaustive or determinative.[61] Moreover, courts consider the factors in relation to one another in order to decide whether a particular use is fair in light of the purposes of copyright.[62] While each factor has its own set of considerations, no one factor is dispositive; they must be all be explored with their results weighed together against the purposes of copyright.[63]

The considerations and implications of each factor are explored below, with particular focus on technological fair uses that depend on the use of copyrighted work for a different (often non-expressive) purpose.

a) Factor 1: "Purpose and Character" of the Use

Under the first factor, purpose and character of the use, courts primarily consider whether the use of the copyrighted work is transformative and also whether the use is commercial or not for profit. Although not clear-cut, many courts utilize a sliding scale approach to evaluate the first factor: the more transformative a work, the less important the commercial aspect becomes.

A transformative use is one in which the purpose of the use is different from the purpose for which the copyrighted work was originally created.[64] Much of the Supreme Court's fair use jurisprudence focuses on transformative uses that supplement the original work with new expression that "provide[s] social benefit by shedding light on an earlier work, and in the process, creates a new one."[65] Subsequent appellate decisions have expanded the scope of transformative uses to include using works in entirely new contexts—privileging uses that supersede the scope of the original

60. 17 U.S.C. § 107 (2006).
61. *Campbell*, 510 U.S. 569 (1994).
62. *Id.* at 578.
63. *Id.*
64. *Id.* at 579.
65. *Id.* at 579.

expression while staying mostly within the context of commentary on the original work.[66] For example, courts have deemed indexing of copyrighted works by search engines to be transformative.[67]

Courts also consider whether the use is commercial or for non-profit purposes under the first factor.[68] The Supreme Court previously proclaimed that commercial use is "presumptively unfair," but the Court has since stepped back and clarified that commercial use does not present "hard presumptive significance."[69] Further, fair use case law shows that the commercial nature of a use is more properly studied under the fourth fair use factor (harm to the market for the copyrighted work).[70]

b) Factor 2: "Nature of the Copyrighted Work"

Under the second fair use factor, courts consider the nature of the underlying copyrighted work; creative works that are closer to the core of copyright should be afforded more protection from copying.[71] Consequently, the scope of protection for creative works such as novels and paintings is larger than works that are mostly functional, such as repair manuals.[72]

In addition to whether the work is creative or functional, courts weighing the second factor may also consider whether the work is out of print. In the Senate Report for the 1976 Copyright Act, the legislature noted that works that are "out of print" and therefore unavailable for purchase through "normal channels" are more likely to qualify for fair use.[73] Following a line of cases that found contradictory rulings based on this report, Congress added a paragraph to the end of § 107 that qualifies the weight of unpublished works: "[t]he 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."[74] This amendment assured that use of unpublished and out of print works would

66. *See* Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007); *see also* Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003).

67. *See id.*

68. 17 U.S.C. 107(1) (2006).

69. *Campbell*, 510 U.S. at 585.

70. Matthew Sag, *Copyright and Copyright and Copy-Reliant Technology*, 103 NW. U. L. REV. 1607, 1646 (2009).

71. 4-13 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05[2][a] (1992).

72. *Id.*

73. *Id.* at § 13.05[2][b][i].

74. Act of Oct. 24, 1992, Pub. L. No. 102-492, 106 Stat. 3415 (current version at 17 U.S.C. § 107 (2006)).

not categorically be deemed fair, thus protecting against instances where the work is made deliberately unavailable.[75]

c) Factor 3: "Amount and Substantiality" of the Use

Under the third factor, courts consider the quantity and quality of the original work that is being used. Consideration of the "amount of substantiality" of the use must be considered "in relation to the copyrighted work as a whole."[76] This analysis does not rest on the substantial similarity between the two works but on whether the amount taken "is reasonable in light of the purpose and the likelihood of market substitution."[77] Consequently, while a review that takes "the heart of the book" may not be fair because it is likely to lessen the demand for the copyrighted work,[78] it may be fair to copy entire photographs when the secondary use only displays small copies in order to connect internet users to websites hosting the image.[79] Similarly, intermediate copying through reverse engineering can be fair use when the entire work is copied in order to uncover the "functional elements of [a software program]" that are embedded within the completed work.[80]

d) Factor 4: The Effect on the "Potential Market for" the Copyrighted Work

The last fair use factor directs courts to consider the use's effect on the "potential market or value of the copyrighted work."[81] According to the Supreme Court, the fourth factor is "undoubtedly the single most important element of fair use."[82] However, just because a potential licensing market

75. 4-13 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05[b][iii] (1992).

76. 17 U.S.C. § 107(3) (2006).

77. Peter Letterese & Assoc. v. World. Inst. of Scientology Enters., 533 F.3d 1287, 1314 (11th Cir. 2008).

78. Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 565 (1985).

79. Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1168 (9th Cir. 2007) (finding that while the amount taken did not favor either party, it was reasonable in light of the minimal potential harm for "the downloading of reduced-size images for cell phones"). *See also* Kelly v. Arriba Soft Corp., 336 F.3d 811, 821 (9th Cir. 2003) (holding that copying an entire photograph is reasonable when the purpose is for connecting users to websites via an image search engine).

80. Sega Enter. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1528 (9th Cir. 1992).

81. 17 U.S.C. § 107(4) (2006).

82. *Harper & Row*, 471 U.S. at 566; *see also* 4-13 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05[4] (1992) ("[I]f one looks at the fair use cases, if not always to their stated rationale, [the fourth factor] emerges as the most important, indeed, central fair use factor.").

could develop does not mean the factor automatically favors the copyright owner.[83] Rather, courts limit the potential licensing revenues by considering only "traditional, reasonable, or likely to be developed markets when examining and assessing a secondary use."[84] For example, in *Campbell v. Acuff-Rose Music, Inc.*, the Court held that a licensing market for "critical reviews or lampoons" was unlikely to develop.[85] Likewise, the *Campbell* Court held that the Copyright Act does not protect against criticism that destroys demand for the original work.[86] Similarly, in *Perfect 10, Inc. v. Amazon.com, Inc.*, the Ninth Circuit rejected plaintiff's claim that Google's transformative search index would damage its market for reduced-size cell phone photographs.[87]

4. Applying Fair Use

a) Transformative Uses in the Technology Sector: Search Engines and Anti-Plagiarism Software

Perfect 10 provides one of the early tests of the fair use doctrine on transformative technologies.[88] At issue in that case was Google's image search engine, which returns thumbnails of images based on their relevance to a user's search query.[89] The lower-resolution thumbnails are transformed into links; when a user clicks on the image, embedded HTML instructions direct the user's Internet browser to the web address where a larger version of the image is republished and stored.[90] Perfect 10 is a business that earns revenue by selling subscriptions to its website that contains copyrighted images of nude models.[91] Perfect 10's website could not be crawled or indexed by Google's image search engine, but it could identify and return results of the images that were stored on third party websites.[92] Because of this, Perfect 10 filed suit against Google for copyright infringement.[93]

The Ninth Circuit held that Google's use of the images as thumbnails was highly transformative because the purpose of the use "served a different function" than artistic expression: it provided a social benefit by

83. American Geophysical Union v. Texaco, 60 F.3d 913, 930 n.17 (2d Cir. 1994).
84. *Id.* at 930.
85. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 592 (1994).
86. *Id.*
87. Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1168 (9th Cir. 2007).
88. *Id. See also* Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003).
89. *Id.* at 1156.
90. *Id.*
91. *Id.* at 1157.
92. *Id.*
93. *Id.* at 1157.

transforming the image into "an electronic reference tool."[94] Moreover, the court held that it did not matter that Google copied the entire image because it was placed "in a new context to serve a different purpose."[95] Lastly, the court rejected the speculative harm to Perfect 10's market for thumbnail cell phone downloads, reasoning that the plaintiff failed to establish that Google users have actually downloaded images onto their cell phones.[96] On balance, the court's fair use analysis focused on the search engine's "significant benefit to the public," finding that its utility outweighed the harms by a speculative market for cell phone downloads.[97]

One year after *Perfect 10*, in *AV ex rel. Vanderhye v. iParadigms, LLC*, two high school students filed copyright suits over the use of their essays and other papers by iParadigms, a company that provides anti-plagiarism software to educational institutions.[98] iParadigms operates the Turnitin software, which provides high schools and universities an automated way to ensure their students' work is original.[99] Instructors submit a body of current and past student work into a database.[100] Once the works are submitted, instructors can check for unaccredited copying by comparing the paper against the database in addition to sources available on the Internet.[101] Like the court in *Perfect 10*, the Fourth Circuit held that the plagiarism detection service was a highly transformative use: academic expression versus plagiarism detection.[102] iParadigms also copied the entire work, but the court found that it was fair considering the purpose, in addition to the fact that the program only displayed portions that enabled users to see the unaccredited copying.[103] Lastly, the court rejected any speculative harm to the market, finding that no market existed for the student papers[104] and that the four factors balanced toward fair use.[105]

As both *Perfect 10* and *iParadigms* demonstrate, courts have used the fair use doctrine to protect uses that enable the public to access and spread

94. *Id.* at 1165.

95. *Id.* (citing Kelly v. Arriba Soft Corp., 366 F.3d at 819; Mattel, Inc. v. Walking Mountain Prods., 353 F.3d 792, 796–98 (9th Cir. 2003)).

96. *Id.* at 1168.

97. *Id.*

98. A.V. ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630, 634 (4th Cir. 2009).

99. *Id.*

100. *Id.*

101. *See id.*; TURN IT IN, http://www.turnitin.com/en_us/products/originalitycheck (last visited February 5, 2013).

102. A.V. ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630, 639 (4th Cir. 2009).

103. *Id.* at 642.

104. *Id.*

105. *Id.*

learning. This preference promotes the overarching Congressional intent of promoting progress and science through knowledge preservation and accessibility. The next Section discusses fair use in the context of two universities making use of the Internet to reproduce copyrighted content for educational purposes.

b) Protection for Educational Reproduction: Streaming Video and E-Reserves

In *Assn. for Info. Media & Equip. v. Regents of the University of California*, the trial court granted defendant's motion to dismiss following an alleged breach of license and subsequent copyright infringement.[106] There, plaintiffs entered into a licensing agreement with University Regents and various individuals within the University of California, Los Angeles ("UCLA") for use of their collection of DVDs.[107] Defendants at UCLA subsequently changed the format of the DVDs in order to make them available via online streaming to students and faculty with access to the UCLA network.[108] Plaintiffs filed suit after UCLA refused to cease the practice.[109] The court granted defendant's motion to dismiss, finding that the agreement allowed UCLA to "publicly perform" the licensed work, not only allowing them to show the work in its DVD format but also to make it available to its community through online streaming.[110] Without much explanation, the court claimed that the making of short term copies was fair use.[111] The court cited *Perfect 10* for this proposition, but UCLA's use was markedly different. Whereas Google displayed thumbnail copies that functioned as a reference tool for internet users,[112] here, the copy was being used for the same primary purpose as the original: viewing its expressive content.

The court also quickly dismissed plaintiff's claim that UCLA's circumvention of protective measures placed on its DVDs was a violation of the DMCA.[113] The court based its dismissal on the fact that plaintiffs could present no more than speculative allegations that UCLA contracted with a third party to bypass the DVD protection system.[114] It is certainly possible

106. Ass'n for Info. Media & Equip. v. Regents of the Univ. of California, CV 10-9378 CBM MANX, 2011 WL 7447148 (C.D. Cal. Oct. 3, 2011) No. CV-10-9378 CBM (MANX), 2011 WL 7447148, at *1 (C.D. Cal. Oct. 3, 2011).

107. *Id.*

108. *Id.*

109. *Id.*

110. *Id.* at *6.

111. *Id.*

112. Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1156 (9th Cir. 2007).

113. *Assn. for Info. Media & Equip.*, 2011 WL 7447148, at *7.

114. *Id.*

that further discovery could have produced more information relating to UCLA's particular use, but the court's hasty dismissal at the 12(b)(6) stage demonstrates the degree to which courts seek to ensure that universities are empowered to increase access to copyrighted works for educational purposes.

Most recently, in *Cambridge University Press v. Becker*, a trial court largely protected Georgia State's use of copyrighted works in its e-reserve system.[115] Three publishers specializing in academic works (backed by the Association of American Publishers and the Copyright Clearance Center) sued Georgia State for its e-reserve policy, alleging that making unlicensed copies of portions of their copyrighted content and distributing it to students infringed their copyrights.[116] Though many of the plaintiffs' claims failed due to their inability to provide a chain of title, the court's fair use analysis found that the purpose of the use, while not transformative, strongly favored George State because it was for teaching purposes at a non-profit educational institution.[117] The court's analysis of the nature of the protected work found that the factor favored the defendants because the works were "intended to inform or educate," making it worthy of less protection than highly creative works such as fictional novels.[118] Analyzing the amount and substantiality of the use, the court disapproved of the university's overly stringent policy for copying but stated that copying small amounts of the works was favored, finding that copying less than ten percent of the underlying work is "decidedly small," and leads toward a finding of fair use.[119] Lastly, the court found that the effect on the market favored the publishers in instances where they could establish a licensing market for digital excerpts.[120] For all other instances, the court held that the factor weighs in favor of Georgia State; without establishing a licensing market, the court found that plaintiffs could not prove that the students would substitute online excerpts by purchasing the complete work.[121] According to Professor James Grimmelmann, this decision, if upheld, would likely lead to the wider exploitation of orphan

115. Cambridge University Press v. Becker, 863 F. Supp. 2d 1190 (N.D. Ga. 2012).

116. *Id.* at 1201.

117. *Id.* at 1225.

118. *Id.* at 1227.

119. *Id.* at 1254. According to James Grimmelmann, this essentially established a bright-line rule that copying less than ten percent of a work guarantees a finding of fair use. *See* James Grimmelmann, *Inside the Georgia State Opinion*, LABORATORIUM, *available at* http://laboratorium.net/archive/2012/05/13/inside_the_georgia_state_opinion (last visited Nov. 11, 2012).

120. *Id.* at 1254.

121. *Id.* at 1256.

works, as there would be no available copyright owner from whom to license (or buy a hard copy from).[122]

As shown above, both congressional and judicial decisions have shown a preference for protecting the uses of copyrighted works for purposes of preservation, access, education, and research.

II. *AUTHORS GUILD V. HATHITRUST*

The second of two suits targeting the Google Books project, *Authors Guild, Inc. v. HathiTrust*, focuses on the defendant's use of digital copies it received in exchange for allowing Google to digitize the works in its collection.[123] HathiTrust placed these copies into the HathiTrust Digital Library ("HathiTrust").[124] The HathiTrust treats works within its collection differently depending on whether the work (1) has a known author, (2) is an orphan work, or (3) is in the public domain.[125]

For in-copyright works with a known author, the HathiTrust stores a digital copy of the work for preservation but also allows its members to conduct full-text searches across the entire collection and allows print-disabled patrons to access the works in the collection.[126] However, while users have the ability to search across the entire collection, the search engine returns only the page numbers on which a particular search term is found in addition to the number of times it appears on the page.[127]

Prior to the suit, the HathiTrust made orphan works fully searchable and available to its users. However, this was only ensuring that the work was not available for sale and that the copyright owner could not be reached—even after publishing the bibliographic information for these orphan works for ninety days.[128] This approach was abandoned following the suit.[129]

The Authors Guild sued HathiTrust for copyright infringement, claiming that (1) the mass digitization violated §§ 108 and 106 of the Copyright Act; (2) the fair use defense was not available to libraries also invoking § 108; (3) injunctive relief was necessary to prevent HathiTrust libraries from making

122. James Grimmelmann, *Inside the Georgia State Opinion*, LABORATORIUM, *available at* http://laboratorium.net/archive/2012/05/13/inside_the_georgia_state_opinion (last visited Nov. 11, 2012).

123. Authors Guild, Inc. v. HathiTrust, No. 11 CV-4351(HB) 2012 WL 4808939, at *1 (S.D.N.Y Oct. 10, 2012).

124. *Id.*

125. *Id.* at *2.

126. *Id.*

127. *Id.*

128. *Id.*

129. *Id.*

their works available to the Google Books project; (4) the planned orphan works project would lead to mass infringement and should therefore be prohibited from continuing; and (5) HathiTrust should be ordered to return all of the unauthorized digital copies within its possession.[130]

The trial court made no ruling on the orphan works project, finding the project unripe for discussion.[131] However, the court quickly disposed of plaintiff's claim that the fair use defense is unavailable to libraries also seeking protection under § 108, holding that the "clear language of Section 108 provides the rights to libraries *in addition* to fair use rights that might be available."[132] Applying § 107 to HathiTrust's use, the court found that both the search indexing and the providing of access to print-disabled individuals are protected by the fair use doctrine.

A. PURPOSE AND NATURE OF THE USE

The court found that the first fair use factor strongly favors HathiTrust because both its search index and access for print-disabled patrons are highly transformative works.[133] Focusing its inquiry on whether the secondary use serves an entirely different purpose, the court held that the search index served an entirely different function than the copyrighted works in that it provided "superior search capabilities."[134] Similarly, the court found that access for printed-disabled users was also transformative because the lack of market for print-disabled individuals demonstrated that their market was not contemplated during the creation of the underlying work, a finding that tended to show the secondary use was substantively different than the original.[135]

B. NATURE OF THE COPYRIGHTED WORK

Although the books at issue were close to "the core of intended copyright protection," the court followed the Second Circuit's instruction in *Bill Graham Archives v. Dorling Kindersley Ltd.* and found that second factor has limited utility when the use is highly transformative.[136]

---

130. *Id.* at *3.
131. *Id.*
132. *Id.* at *8.
133. *Id.* at *11.
134. *Id.*
135. *Id.* The cursory treatment of this function is more thoroughly analyzed in light of the use's effect on harm to potential markets.
136. *Id.* at *12 (citing Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 612 (2d Cir. 2006) ("the second factor may be of limited usefulness where the creative work is being used for a transformative purpose")).

C. AMOUNT OF THE WORK COPIED

As with the second factor, the court found that the third fair use factor, the amount and substantiality of the use of the original, did not weigh in favor of either party due to the transformative nature of HathiTrust's use.[137] While the entire work was copied, the court considered whether the extent of the copying "was reasonable in relation to the purpose."[138] Finding that copying the entire work was necessary to make the search index and to ensure print-disabled accessibility, the court held that the factor did not favor the Authors Guild.[139]

D. EFFECT ON THE MARKET FOR OR VALUE OF THE WORKS

Due to the non-commercial nature of HathiTrust's use, the court found that the Author's Guild was required to show "by a preponderance of the evidence that some meaningful likelihood of future harm exists."[140] First, the court found that purchasing additional copies would not have allowed HathiTrust to engage in its transformative use and that the plaintiff's claim that HathiTrust's security measures were inadequate was unsupported."[141] Second, discussing the harm to a potential licensing market, the court limited the analysis to "traditional, reasonable or likely to be developed markets."[142] The court found that, while a collective licensing market could develop through the Copyright Clearance Center, the law did not require it to consider potential markets.[143] The court also found that a licensing market to provide print-disabled individuals access to works on the scale of the HathiTrust project was unlikely[144] and pointed to Congressional efforts to protect print-disabled individuals via the Americans with Disabilities Act.[145]

On balance, the court concluded that the fair use factors suggested that enhanced search capabilities, the preservation of fragile works, and the enablement of print-disabled individuals to access the collection's wide body of work all support copyright's goal of promoting the useful arts and

137. *Id.*
138. *Id.* (citing Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449–50 (1984)).
139. *Id.*
140. *Id.* at *13 (citing *Sony*, 464 U.S. at 451).
141. *Id.*
142. *Id.* at *13 (citing American Geophysical Union v. Texaco, Inc., 60 F.3d 913, 930 (2d Cir. 1994)).
143. *Id.* at *14 (citing Sony, 446 U.S. at 456).
144. *Id.*
145. *Id.*

sciences.[146] The court held that, in comparison to the large social utility promised by the uses, "the authors stand to gain very little if the public is deprived of this resource."[147]

Lastly, the trial court held that HathiTrust's reproduction of "published non-dramatic works" for print-disabled individuals was protected by the Chafee Amendment to the Copyright Act, which protects the reproduction and distribution of copies of non-dramatic works "in specialized formats exclusively for use by the blind or other persons with disabilities."[148] While the exemption is only available to an "authorized entity," which is a non-profit organization or governmental agency with a "primary mission to provide specialized services" for disabled individuals, the court found that HathiTrust qualifies for the exemption because it has a goal of providing access for print-disabled individuals.[149]

III. THE AFTERMATH OF *HATHITRUST V. AUTHORS GUILD*: WHAT NOW?

The *HathiTrust* decision solidifies a growing judicial commitment to protecting libraries and educational instructions in their efforts to make use of technology to increase preservation efforts, modernize their pedagogy, and facilitate better research. Combined with earlier judicial efforts to protect web crawling and image search tools in order to connect the public with the vast repository of information available on the Internet, courts are slowly removing barriers to using the technological advancements of the Digital Age.

However, these new judicial protections are not free from problems. In an attempt to protect useful technologies, courts have strained the language of various statutes to allow projects to move forward. Stretching existing laws to achieve favorable results, although seemingly helpful, leaves opinions vulnerable on appeal. Moreover, the results reached in these cases may be better implemented by Congressional redrafting. But given the slim possibility that Congress will take action, the judiciary's patches become increasingly important as they signal to competing interests the ways in which they must coexist.

146. *Id.*

147. *Id.* at *15 (quoting Library Amici Brief, Authors Guild, Inc. v. HathiTrust, No. 11 CV 6351 HB, 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012)).

148. *Id.* (citing 17 U.S.C. § 121 (2006)).

149. *Id.*

Unfortunately, many of the decisions regarding the use of technology for educational and archival purposes are still in their infancy. Most of these cases are certain to face appeals, and it may be several years before we have judicial opinions that carry sufficient weight to settle questions regarding mass digitization and copy-dependent technology. The Authors Guild filed its notice of appeal in the *HathiTrust* decision less than a month later after the trial court's opinion, and the opinion faces some difficult questions that may narrow the scope of the fair use ruling.[150] First, this Section discusses the future of the *HathiTrust* decision on appeal and the decision's potential effect on the Authors Guild's case against Google. The Section argues that both uses should be found fair because they are transformative works with a useful public benefit that outweighs the harm to the Authors Guild. Given the growing impetus towards protecting technologies that enable preservation and access, an appellate decision from the Second Circuit is desirable because it will largely settle the law in the wake of Congressional inaction. Lastly, this Section discusses updates to the Copyright Act that would enable competing interests to coexist in the Digital Age.

1. HathiTrust on Appeal: Protecting the Search Index Brings the Second Circuit in Line with the Third Wave of "Transformative" Uses

More than twenty years have passed since Judge Leval's "Toward a Fair Use Standard" emphasized the purpose and nature of a secondary work as "the soul of fair use."[151] According to Judge Leval, transformative uses should be favored because they promote the creation of new works and encourage interaction with the works that came before them—"[adding] value to the original."[152] These active engagements promoted the progress of science and the useful arts by using the primary work "as raw material" and transforming it through the "the creation of new information, new aesthetics, new insights and understandings" that the fair use doctrine must protect "for the enrichment of society."[153] Following the publication of Judge Leval's article, courts around the country began taking on a more expansive view of "transformative" and employed the idea to protect new technological uses of copyrighted works that provided a public benefit.

150. Gary Price, *Authors Guild Appeals HathiTrust Decision, Library Copyright Alliance Issues Statement*, INFODOCKET (Feb. 2, 2013, 11:37 PM), http://www.infodocket.com/2012/11/09/authors-guild-appeals-hathitrust-decision-library-copyright-alliance-issues-statement.

151. Judge Pierre N. Leval, *Toward A Fair Use Standard*, 103 HARV. L. REV. 1105 (1990).

152. *Id.* at 1111.

153. *Id.*

The first wave of these decisions sought to protect intermediate copying for purposes of reverse engineering. For example, in *Sega Enterprises Ltd. v. Accolade, Inc.*, the Ninth Circuit protected Accolade's reverse engineering of Sega's source code in order to create video game cartridges that would run on Sega's Genesis player.[154] There, Accolade was seeking to learn the "functional" aspects of the Sega Genesis so that it could take that information and create its own original video games that were compatible with Sega's console.[155] While there was no commentary or engagement with the previous work, implicit in the court's rationale was the feeling that society would be well-served by not allowing Sega to monopolize the market for both the console and the games that ran on it.

The second wave of cases protected secondary works that made slight alterations to copyrighted works and incorporated them into their secondary uses. In *Bill Graham Archives v. Dorling Kindersley, Ltd.*, the Second Circuit held that the fair use doctrine protected the incorporation of a copyrighted poster into a new work depicting the famous 1960s band, the Grateful Dead.[156] There, the defendant created a 480-page coffee table book that told the story of the Grateful Dead along a timeline that ran throughout the whole work, which contained reduced-size prints of the posters and tickets along with text and other graphic art.[157] The court found that the purpose of the concert posters was to provide information regarding upcoming concerts, whereas the book used the posters and tickets as a historical artifact in its documentation of the bands history.[158] Unlike the *Sega* line of decisions, where the challenged uses competed for the same market, the *Bill Graham* wave of decisions protected works that used the underlying work for a different purpose and competed in different markets.

Most recently, a third wave of fair use decisions have favored transformative technologies that use copyrighted works for a different purpose that ultimately provides a useful public benefit. Beginning with *Kelly v. Arriba Soft Corp.* and continuing with *Perfect 10, Inc. v. Amazon.com, Inc.*, the Ninth Circuit protected two online service providers that indexed thumbnails of copyrighted works and transformed them into pointers that connected

154. Sega Enter. Ltd. V. Accolade, Inc., 977 F. 2d 1510, 1527 (9th Cir. 1992). *See also* Sony Computer Entm't v. Connetix Corp., 203 F.3d 596, 602 (9th Cir. 2000) (protecting reverse engineering that made PlayStation games compatible with PC operating systems).

155. *Id.* at 1523.

156. Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 615 (2d Cir. 2006).

157. *Id.* at 607.

158. *Id.* at 609.

users with copyrighted content available on the Internet via image search.[159] In both cases, the purposes were commercial and contained no commentary along with the original work.[160] Nonetheless, the indexing and resulting search capabilities provided a substantial public service by connecting users with the vast amount of content available on the Internet.[161] As discussed in Part I, promoting access to copyrighted works has long been a goal of copyright law, and on balance, the technologies promoted the progress of science and the useful arts by enabling the public to find a plethora of "raw material" to use in the creation of new works. Subsequent to the *Perfect 10* decision, the Fourth Circuit also protected technology that used copyrighted works in its plagiarism detection software because it did not seek to display its expressive content but rather to cross-check works for originality.[162]

HathiTrust's searchable index of library collections provides a substantial public benefit, as it provides an accessible database that far eclipses the search-ability of any existing library index. Moreover, with the inevitable deterioration facing aging collections, HathiTrust's digital collection ensures that large-scale preservation efforts can be successful. Lastly, by making its collection available to print-disabled individuals, it empowers an underrepresented segment of the population to access thousands of books that would have otherwise remained unavailable.

Judge Baer's decision in *HathiTrust* mentioned each of these benefits, signaling the extent to which the court aligned itself with the decisions from the Ninth and Fourth Circuits. The Authors Guild's appeal provides the Second Circuit the opportunity to create clear case law on whether the fair use doctrine protects technologies that use copyrighted works to connect users with information instead of expressive content. Indeed, the trial court quoted *Bill Graham*, *Perfect 10*, and *iParadigms* in its fair use analyses, embracing the more expansive interpretation of "transformative use."[163] The trial court held that HathiTrust's database served "an entirely different purpose than the original works: the purpose is superior search capabilities rather than actual access to copyrighted material."[164] The search capabilities

159. *See* Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007); Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003).

160. *Id.*

161. *Id.*

162. *See infra* Section I.B.4.

163. Authors Guild, Inc. v. HathiTrust, No. 11 CV-4351(HB) 2012 WL 4808939, at *11 (S.D.N.Y Oct. 10, 2012).

164. *Id.* at *12.

of the HathiTrust database have already given rise to new methods of academic inquiry, such as text mining.[165]

The trial court's fair use opinion will likely be upheld on appeal. Judge Baer's opinion did a thorough job grounding the fair use finding in the existing body of jurisprudence inside and outside the Second Circuit. His ruling reflects the growing sentiment that there is something notably different that motivates the unauthorized copying behind non-expressive use that should place it outside the scope of copyright protection. Indeed, as Professor Matthew Sag said, "once the pivotal nature of expressive substitution is properly understood, the implications for copy reliant technologies crystallize: the non-expressive use of a copyright work should not ordinarily result in copyright infringement."[166] However, not all aspects of the trial courts opinion are free of problems. In that vein, the next Section discusses the fair use ruling regarding copies for the print-disabled.

2. The Trial Court's Fair Use Protection for Print-Disabled Individuals May Be Vulnerable on Appeal

Applying the same logic that protected HathiTrust's search index, the trial court also held that making copies for the print-disabled was fair.[167] Whereas much of the rationale regarding transformative use applied to the search index, it is not so readily applicable to copying for the print-disabled. Unlike the index—which only displayed the pages on which a search term appeared—the print-disabled copies reproduced the copyrighted works in their entirety. Consequently, while transformative, the print-disabled copies still reproduced the original works' expressive content—a right reserved to copyright owners under their derivative work right.[168]

As if anticipating objection, the court cited the Americans with Disabilities Act ("ADA") privileges under § 121.[169] Unfortunately, that analysis is also vulnerable, as it depends on an expansive interpretation of who qualifies as a "protected entity" under the statute. Section 121 states that protection is conditioned on being a non-profit or governmental organization whose "primary mission" is to provide "specialized services relating to training, education, or adaptive reading or information access needs of blind or other persons with disabilities."[170] Moreover, the statute

165. *Id.* at *14.

166. Matthew Sag, *supra* note 70, at 1638 (2009).

167. Authors Guild, Inc. v. HathiTrust, No. 11 CV-4351(HB) 2012 WL 4808939 (S.D.N.Y Oct. 10, 2012).

168. *See* 17 U.S.C. § 106(2) (2006).

169. *HathiTrust*, 2012 WL 4808939, at *15.

170. 17 U.S.C § 121.

only protects the reproduction of "nondramatic literary works" intended as training and vocational material.[171] Given that "approximately 76 percent of the identified works are fiction," it is unclear whether HathiTrust's copies qualify for protection under § 121.

Nonetheless, there are strong policy considerations that justify upholding Judge Baer's ruling. First, as noted by the trial court opinion, the print-disabled community is dramatically under-served. Less than five percent of copyrighted works are accessible for the print-disabled, and a market to make these works available via licensing has not emerged.[172] This may be the result of low market demand; according to the University of Michigan, only thirty-two students had signed up to access its print-disabled copies.[173] It is likely that a significant market may never develop, and the Second Circuit should protect HathiTrust's use given the motivations behind the ADA. However, it is also possible a judge will read the statute strictly to hold that it is not the judiciary's role to expand the scope of protection that Congress provided. Nonetheless, given no threat of market substitution and copyright's goal of promoting creativity and learning, § 121 should protect the print-disabled community and allow them to access the enormous body of works that have remained unavailable to them for decades; a narrow ruling may finally prompt Congress to revise the statute.

3. Same Song, Different Tune: Implications for the Case Against Google Books

Google Books takes the scans of millions of books held in library collections and makes them fully searchable by anyone with an Internet connection.[174] The product also provides links to where a book may be borrowed from a library or purchased from a bookseller.[175] The corpus contains everything from fiction and nonfiction to reference books and cookbooks.[176] The product is available for free, but Google places ads next to the search results.[177]

171. *Id.*

172. *HathiTrust*, 2012 WL 4808939, at *14.

173. *Id.*

174. *See Google Books Library Project*, ABOUT GOOGLE BOOKS, http://www.google.com/googlebooks/library.html (last visited Mar. 28, 2013).

175. *Id.*

176. Def.'s Mem. Supp. Mot. for Summ. J. at 7, Authors Guild, Inc. v. Google Inc., 770 F. Supp. 2d 666 (S.D.N.Y. 2011) (No. 05 CV 8136(DC)).

177. *Id.* at 8.

When a user enters a query, the program returns a list of books where the search term appears.[178] In some situations, users will also see snippets from the book to provide context for where the term appears.[179] Google places security measures so that the book cannot be reconstructed, and the results are presented as an image, so they cannot be copied and pasted.[180] Public domain works are fully viewable on Google Books, and publishers may request that certain pages be viewable for advertising purposes.[181] Indeed, many publishers already do this because of the significant marketing benefit.[182]

Applying *HathiTrust*'s fair use analysis to Google Books yields similar results. Google Books's use is highly transformative, and its commercial nature should be of limited importance given the extent to which it promotes copyright's goal of increasing preservation and access in order to promote the progress of science and the useful arts.

First, Google Books does not display copyrighted content for its artistic expression but to connect users to works that are relevant to their queries. In many ways, Google Books represents a card catalog for the Digital Age. But rather than being limited by the categories chosen by librarians, Google Books empowers users to search for specific words and phrases and across millions of seemingly unrelated books. Admittedly, Google Books goes beyond a card catalog by displaying the text of the work in addition to information about it, but only to provide context for the location where a term appears; the snippets never reveal enough context to allow a user to substitute for purchasing the work or borrowing a copy from a library.[183]

Second, Google Books contains millions of works with varying degrees of copyright protection. Some works, such as the challenged fictional works, are closer to the core of copyright and entitled the broader protection.[184] But others, such as academic works (which make up a majority of the corpus), instruction manuals, and dictionaries, have a thinner scope of protection because of the underlying facts they creatively express and arrange.[185]

178. *Id.*
179. *Id.* at 1.
180. *Id.* at 10.
181. *Id.* at 11.
182. *Id.* at 17.
183. *Id.* at 10.
184. *See* Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 586 (1994).
185. Indeed, a partnership of professors has filed an amicus brief objecting to the class certification sought by the Authors Guild, arguing that the class inadequately represents the interest of academic writers that wish to have Google Books upheld as a fair use. *See* Brief of Amici Curiae Academic Authors in Support of Defendant-Appellant and Reversal, *available*

Nonetheless, as the court held in *HathiTrust*, the nature of the copyrighted works is of limited importance when dealing with highly transformative works.[186]

Third, the search functionality only provides useful results if the entire work is scanned. However, rather than minimizing the importance of this factor, challenged uses should be analyzed to ensure that the amount and substantiality used is fair, especially if the secondary use is commercial. Here, Google must use the entire copyrighted work for its search index. This use is not expressive because the product is not designed to display the copyrighted work. However, the search results contain snippets of the work in order to provide users with the context in which their search terms appear. The amount of the work that gets displayed should be analyzed in order to ensure it does not reproduce enough of the work to produce harm to the copyright owner. As in *Cambridge*, where the court held that the amounts of copyrighted articles displayed were reasonable for its educational purpose, here Google Books displays only enough to provide context and blocks out enough of an overall work to ensure that its expressive content cannot be recreated. In light of Google's research purpose, the amount and substantiality taken and displayed is fair.

Fourth, Google Books enhances the market for copyrighted works by connecting readers with books relevant to their interests and queries. By providing users with links to where they can obtain the works, Google Books expands the demand for books that could otherwise remain obscure. Whereas HathiTrust did not harm the market because it was only available to its library patrons, Google Books expands the market for the underlying works in its corpus. Google Books provides a service that combines modern search with the casual browsing that typified thumbing through a book at a bookstore.

Neither HathiTrust nor Google engage in a direct commercialization of copyrighted content. As the Supreme Court held in *Sony Corp. of America v. Universal City Studios, Inc.*, a use that has no demonstrable effect upon the potential market for or the value of the copyrighted work need not be prohibited in order to protect the author's incentive to innovate.[187] The potential loss from licensing revenue should not prevent the Google Books project from moving forward because the speculative market does not exist, and the product provides a significant public benefit that enables wide-scale

at http://thepublicindex.org/docs/cases/authorsguild-2ndcir/32-academic-amicus.pdf (last visited Feb. 5, 2013).

186. *HathiTrust*, 2012 WL 4808939, at *12.

187. Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 450 (1984).

preservation and access. As the Second Circuit held in *Bill Graham*, "were a court automatically to conclude in every case that potential licensing revenues were impermissibly impaired simply because the secondary user did not pay a fee for the right to engage in the use, the fourth factor would always favor the copyright owner."[188] Instead, the harm to potential licensing revenues should be measured by balancing the public benefit with the negative effect it may have on authors' incentive to create new works.

Balancing the four factors, Google Books's fair use defense is likely as strong as HathiTrust's. While Google's service goes beyond the preservation and research purposes of HathiTrust, it provides access to the public, enabling large-scale preservation, access, and search-ability that is unparalleled (even by the Library of Congress). Balancing the public benefit against the potential loss of revenue, protecting Google Books does far more to protect the promotion of the arts and useful sciences. Given the current inability of Congress to provide meaningful copyright reform for technologies with highly transformative purposes that facilitate the goals of § 107's preamble, fair use is the best place for judicial bodies to protect useful technologies that pose relatively small threats to the incentive for innovation.

IV. CONGRESSIONAL REFORM, WHILE IDEAL, IS UNLIKELY

Congressional updates to the Copyright Act, although an unlikely solution, could ensure that copyright law is compatible with the technological advancements of the Digital Age. First, this Part suggests that Congress consider updating the statutory exemptions available to libraries and research institutions under § 108 of the Copyright Act to enable increased access, ensure preservation, and enable research. Second, this Part suggests that Congress expand the number and scope of "protected entities" under § 121 of the Copyright Act. Lastly, this Part suggests that Congress consider categorically removing non-expressive uses from the realm of copyright.

A. EXPANDING § 108 FOR LIBRARIES AND EDUCATIONAL INSTITUTIONS

Just as the original exemptions for libraries under § 108 grew out of the "photocopying revolution," the Internet Age has dramatically expanded the scope of reproduction and distributions due to new technology, often resulting in a sharp decrease in cost.[189] The challenges to book digitization and college e-reserves are just two examples of the impediments caused by

188. Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 614 (2d Cir. 2006).
189. *See* Matthew Sag, *supra* note 70, at 1612.

the current law's inability to properly keep up with the pace at which libraries and universities are modernizing their practices.

Rather than allowing the effect of isolated court decisions to slowly trickle down, a reformation of the Copyright Act should be completed. Section 108 privileges for libraries is a logical point at which to begin this modernization effort because Congress has already allowed libraries to engage in non-commercial photocopying without fear of direct or indirect liability for copyright infringement.[190] As before, the Copyright Office should issue a study that obtains input from libraries, educational institutions, technology companies, and copyright holders to determine the extent and scope of any changes to § 108.

First, the modernization of § 108 should expand the scope of protected entities from "libraries and archives" to also include "educational institutions." Each of these entities should be privileged in their use of digital copies for expressive and non-expressive uses for purposes of research, access, education, and preservation. Just as the existing § 108 dealt with the concern that technological advances would eclipse the market for their works, the expanded § 108 would require strict compliance with statutory provisions in order to obtain protection. Moreover, the statutory requirements could be supplemented to ensure library and educational copying is limited to the amount necessary for its purpose. A statutory update would be a better place for a bright line rule similar to the one mentioned by Judge Evans in *Campbell*.[191] This would ensure that the amount used properly reflects the underlying needs without requiring librarians and educators to engage in speculative fair use analyses that even well-versed judges cannot agree on.

As a tradeoff for the restrictions on use, the expanded § 108 should also exempt transformative uses that require copying the entire work. This protection is necessary to ensure the projects such as the HathiTrust search index retain social utility for the public. Under the aegis of § 108, institutions seeking to enable users to query across their entire collections would not have to implement an opt-out mechanism that compromises their ability to provide a truly useful and comprehensive search index. This allows public and private universities to continue Congress's goal of increasing preservation and access in order to promote the progress of science. The benefit available to the public substantially outweighs the potential harm to

190. *See* Copyright Act of 1976, § 108, 90 Stat. 2541.

191. Cambridge Univ. Press v. Becker, 863 F. Supp. 2d 1190, 1236 (N.D. Ga. 2012) (holding that using less than 10% of a copyrighted work is likely to constitute fair use for educational purposes).

the market for the books because the search functionality is not an expressive substitution for the creative expression in the works.

B. EXPANSION OF § 121 PROTECTION

Section 121 should be expanded to allow print-disabled individuals to access the large volumes of works that have remained inaccessible to them for decades. Reproduction of copyrighted works for individuals with print-disabilities is made easier by the advent of digital copies. As noted by the court in *HathiTrust*, there is no market for providing print-disabled individuals access to the vast body of works that are housed in our nation's libraries.[192] Unfortunately, the language of § 121 only protects "authorized entities," which are nonprofit organizations or governmental agencies with a "primary mission" to provide print-disabled individuals with "specialized services relating to training, education, or adaptive reading or information access."[193] It is difficult to imagine an interpretation of the purposes of copyright law that would not promote connecting every member of the public with the vast and expansive body of works. Given that there has been no attempt to establish a commercial market, § 121 should be expanded to protect the reproduction and distribution of copyrighted content for print-disabled individuals. Moreover, the meaning of "authorized entities" should be expanded to include all non-commercial uses. This would enable Google Books to also make its corpus available to the print-disabled minority, albeit a special version that does not contain ads. As stated by Judge Baer,[194] the absence of a market is proof that the benefit of this service cannot be outweighed by a potential loss to the copyright owner. Indeed, the authors have allowed the loss to occur for decades.

V. CONCLUSION

Digital libraries provide the promise of superior search capabilities and a more effective way to connect readers around the world with the places where they can purchase and borrow books. This technical breakthrough decentralizes the collection of knowledge from university archives, allowing all users to have access. While neither Google Books nor HathiTrust provides a digital public library, each makes use of mass digitization in a way that advances the purposes of copyright by facilitating large-scale preservation and enabling the public to access a wide range of material. The

192. Authors Guild, Inc. v. HathiTrust, 11 CV 6351 HB, 2012 WL 4808939, at *15 (S.D.N.Y. Oct. 10, 2012).

193. 17 U.S.C. § 121(d)(1) (2006).

194. *HathiTrust*, 2012 WL 4808939, at *15.

concern for access and preservation informed all iterations of the Copyright Act, and its evolution has consistently sought to be for the benefit of the public. Education and research provide the means by which authors can create new works of science and useful arts; technologies that facilitate these goals without expressing the underlying content of copyrighted works should be protected.

Emerging technology will continue to challenge copyright law. Consequently, updates to the Copyright Act will periodically prove necessary. The fair use doctrine provides the flexibility necessary to weigh the benefits of disruptive technologies against the importance of securing limited monopolies to the authors of copyrighted work. The *HathiTrust* decision solidifies a growing preference for educational and research institutions making use of technologies reliant on copying to facilitate learning.

However, for all the promise of emerging technology, judicial reaction to the litigation signals to competing interests the degree to which they must coexist. The Google Books project represents a product that was designed cognizant of incoming lawsuits, and its design is one that sought to balance rights holder concerns by displaying no more than is necessary for user queries and installing security measures that prevent the product from substituting the demand for copyrighted content. This type of behavior must be encouraged, as the scope of transformative use must be balanced by the incentive provided to authors by providing them with exclusive rights regarding reproduction. It is that balance that will ensure that copyright law continues to foster the creation and dissemination of original work.