

# GOOGLE'S TECHNICOLOR DREAMCOAT: A COPYRIGHT ANALYSIS OF THE GOOGLE BOOK SEARCH LIBRARY PROJECT

*By Emily Anne Proskine*

I am certainly not an advocate for frequent changes in laws and constitutions. But laws and institutions must go hand in hand with the progress of the human mind. As that becomes more developed, more enlightened, as new discoveries are made, new truths discovered and manners and opinions change, with the change of circumstances, institutions must advance also to keep pace with the times. We might as well require a man to wear still the coat which fitted him when a boy . . . .<sup>1</sup>

Bringing copyright law up to speed in an era of advancing digital technology is exactly one of those events along the course of history to which Thomas Jefferson referred above. New discoveries, truths, and opinions brought about by the digital age require institutions to advance and legislation to be updated. Lagging behind innovations in technology, the coat of copyright law is getting a little too tight. Thus, law must advance to keep pace with the times. However, the required alterations are not clear. Will it suffice to let out the hem of copyright's coat, or must one be tailor-made from scratch in order to bring copyright law into the current era?

The Google Print Library Project, renamed the Google Book Search Library Project<sup>2</sup> ("Google Library Project"), more so than any of the other

---

© Emily Anne Proskine

The author hereby permits the reproduction of this Note subject to the Creative Commons Attribution-ShareAlike 2.5 License, the full terms of which can be accessed at <http://creativecommons.org/licenses/by-sa/2.5/legalcode>, and provided that the following notice be preserved: "This Note was first published by the Regents of the University of California in the Berkeley Technology Law Journal's Annual Review of Law and Technology."

1. Inscription on the Thomas Jefferson Memorial in Washington, D.C., taken from a letter from Thomas Jefferson to Samuel Kercheval (July 12, 1816), *available at* <http://www.nps.gov/thje/memorial/inscript.htm>.

2. Google rebranded Google Print, which includes the Library Project, as the Google Book Search, in order to better describe the project's purpose. *See Google Print Renamed Google Book Search*, MARKETINGVOX, Nov. 18, 2005, *available at* [http://www.marketingvox.com/archives/2005/11/18/google\\_print\\_renamed\\_google\\_book\\_search](http://www.marketingvox.com/archives/2005/11/18/google_print_renamed_google_book_search).

library digitization projects that have emerged over the last six months,<sup>3</sup> presents the perfect opportunity for an analysis of the “fit” of current copyright law in the age of the internet. With unprecedented scope and huge corporate exposure, the Google Library Project represents a foray into uncharted waters of intellectual property law.

Part I of this Note will evaluate the Google Library Project, examining its scope, plan for implementation, competitors, cost, and potential copyright implications. Part II of this Note will address the possible legal repercussions for Google and assess the probability of Google’s success moving forward under the current copyright regime. Specifically, this Part will address the copyright issues implicated by the Library Project, namely the adherence of copyright liability and the applicability of the fair use defense. Part III will present a legislative patchwork solution to the current copyright regime that would allow socially beneficial projects like Google’s, which fail to harm copyright owners or impede their immediate or potential markets, to proceed. Part IV will address a new copyright regime, one that is based on a distribution right and tailored specifically to address the problem of unauthorized distribution of copyrighted works in the digital era.

## I. THE GOOGLE LIBRARY PROJECT

In the digital age, it is not surprising that one of the biggest players in the digital market would start looking at new and improved venues for information organization and access. While Google currently indexes approximately eight to ten billion items online,<sup>4</sup> the quality of the content provided by many of those sources is questionable. Internet resources are unknown, content is fluid, and authors are unreliable because just about

---

3. Yahoo! and MSNSearch announced their public domain digitization project in October 2005. See Stefanie Olsen, *An Open-Source Rival to Google’s Book Project*, CNET NEWS.COM, Oct. 26, 2005, available at [http://news.com.com/an+open-source+rival+to+googles+book+project/2100-1025\\_3-5915690.html](http://news.com.com/an+open-source+rival+to+googles+book+project/2100-1025_3-5915690.html). Amazon.com announced its new Amazon Pages program in November 2005. See *Amazon Takes Page from Apple*, SEATTLE TIMES, Nov. 5, 2005, at D1. Random House, the world’s largest publisher of trade books, was the first big publishing house to move into online book access in November 2005. See *Pulp Friction*, ECONOMIST, Nov. 12, 2005, at 83. Amazon is working on a model of pay-per-view charging as well. See *id.*

4. *Id.*

anyone can anonymously create a website and post it on the internet at little or no cost.<sup>5</sup>

Additionally, studies by Microsoft have shown that around half of all search queries performed on the internet fail to provide the information that users want.<sup>6</sup> This is at odds with the accessibility and searchability the internet promises. By contrast, physical libraries house print books whose trusted contents have been scrutinized by their authors, revised by their editors and ultimately approved by their publishers. Each book is the result of a laborious process of research, redaction, and outlays of time, energy and money.<sup>7</sup> Nonetheless, the physical nature of print books and brick-and-mortar libraries has posed a much greater challenge to users in search of the knowledge within. Print books are difficult to locate, transport, and share; their physical nature exposes them to deterioration; and they are expensive. Accordingly, studies show patronage at libraries nationwide is on the decline.<sup>8</sup>

#### A. A Brief History of Digital Library Projects

More than three decades ago, the intersection of the print world and the computer, with its extraordinary ability to search and store information, generated several digital library initiatives. Project Gutenberg began in the 1970s, and now counts more than 17,000 e-books in nearly 45 languages in its digital collection.<sup>9</sup> The Internet Archive is another, more re-

5. See Christopher Heng, *How to Design and Publish Your Website*, THE SITE WIZARD.COM, <http://www.thesitewizard.com/gettingstarted/mozillacomposer1.shtml> (last visited Mar. 13, 2006).

6. *Pulp Friction*, *supra* note 3, at 83.

7. Only about 20% of a publisher's budget for each book pays for paper, printing, and binding. The rest pays for the publisher's overhead, cuts taken by distributors, booksellers, and promoters. The author's cut is quite low. As a general rule it ranges from 10% to 15%. And prices for books are increasing. From 1975 to 2000 the price of the average hardcover book of fiction went up 200%, poetry and drama 211%, and nonfiction 123%. Christopher Dreher, *Why Do Books Cost So Much?*, SALON.COM, Dec. 3, 2002, <http://www.salon.com/books/feature/2002/12/03/prices/print.html>.

8. Library visits in 1998 were down 9.2% from 1994. ALA, *Academic Libraries in the U.S.: Statistical Trends*, <http://www.ala.org/ala/ors/statsaboutlib/academiclibraries.htm> (last visited Mar. 13, 2006).

9. Michael S. Hart, *History and Philosophy of Project Gutenberg*, PROJECT GUTENBERG, Aug. 1992, <http://www.gutenberg.org/about/history>. Project Gutenberg began in 1971 when Michael Hart was given an operator's account with \$100,000,000 of computer time by the operators of the Xerox Sigma V mainframe at the Materials Research Lab at the University of Illinois. Hart concluded that the most valuable manner in which to use his time would be to make use of the storage, retrieval and searching capabilities of the computer. Starting with the Declaration of Independence, Hart began

cent initiative, a nonprofit founded to build a digital library offering permanent access for researchers, historians, and scholars to historical collections that exist in digital format.<sup>10</sup> Founded in 1996, the Internet Archive receives data donations and has grown to include texts, audio, moving images, software and archived webpages in its collections.<sup>11</sup> In the passing years, most large university libraries in the United States began digitizing their rare or unique collections, generally for archival purposes.<sup>12</sup> More recently, companies and organizations have also begun digitizing new and old print materials for easier storage and access online. Until now, however, inadequate technical capabilities and the high costs of such projects have limited digitization efforts.

## B. The Scope of the Google Library Project

In response to these obstacles and growing consumer demand for improved search engine results, Google announced its Library Project on December 14, 2004.<sup>13</sup> The goal of the project is “to make the full text of all the world’s books searchable by anyone” with a computer and internet access.<sup>14</sup>

One thing is certain: the project is ambitious. Google ultimately plans to include in its database digital copies of every work in the collections of “the Google 5,” the name given to the five participating library collections of Stanford University, the University of Michigan, Harvard University, Oxford University, and the New York Public Library.<sup>15</sup> To incentivize library participation, Google will set up a simple exchange—for each book

---

reproducing texts onto his computer. Today, the Project focuses on providing access to public domain or donated works that are of great public interest. *Id.*

10. See Internet Archive, About IA, <http://www.archive.org/about/about.php> (last visited Mar. 13, 2006).

11. *Id.*

12. See UM Library/Google Digitization Partnership FAQ, Aug. 2005, at 2, <http://www.lib.umich.edu/staff/google/public/faq.pdf> [hereinafter UM Library/Google FAQ].

13. See Press Release, Google, Google Checks Out Library Books (Dec. 14, 2004), <http://www.google.com/press/pressrel/print-library.html>.

14. Posting of Adam M. Smith, Google Print Project Manager, to Google Blog (Aug. 11, 2005, 13:31 PDT), <http://googleblog.blogspot.com/2005/08/making-books-easier-to-find.html>.

15. To start, Google will scan and digitize all eight million of Stanford University’s print volumes, all seven million volumes at the University of Michigan, a pilot of 40,000 randomly selected books of Harvard University’s 1.5 million volumes, a sample of the New York Public Library’s twenty million items, and one million public domain volumes at Oxford’s Bodleian Library. Deborah Lines Andersen, *Benchmarks: The Google Library*, 7 J. ASS’N FOR HIST. & COMPUTING 3 (2004).

a participating library gives to Google, Google will provide to that library a digital copy of the book. Google will make these books searchable on its search engine anywhere in the world, and at no cost.<sup>16</sup> What the participating libraries will do with their digital copy is unclear. Nonetheless, over the course of the next five to seven years, Google plans to scan and digitize over fifteen million books in the initial stages of this project.<sup>17</sup>

### C. Google's Plan for Implementation

In order for this large-scale project to function efficiently, Google must scan and digitize the full text of all books in the Google 5 collection in order to “unbundle” the content from the printed page. Google has not made public its scanning technology, citing the proprietary nature of the hardware and software.<sup>18</sup> However, the technology may closely resemble the specialized scanning machine and the open-source software that the Internet Archive built specifically for the purpose of scanning and digitizing books: “The ‘machine’ is an assembly of a standard PC with the software installed, two digital cameras, a pedal-operated glass and metal stand to hold and secure books at an angle, along with a table and a chair.”<sup>19</sup> The “machine” is operated by one person who manually flips the book’s pages by hand, photographing each page.<sup>20</sup> The pages of the book then appear on the computer screen in their original form.<sup>21</sup> The operator enters some metadata about the book—the author, title and publication date.<sup>22</sup>

The digitized contents of the book can then be searched by Google’s search engine, which crawls the content and ultimately presents accurate results in response to a user query.<sup>23</sup> While this necessitates that every page of every book in the Google Library collection would be searchable

---

16. *Id.* This content is already being scanned and is now searchable on the Google Book Search Website at <http://books.google.com> (last visited Mar. 13, 2006).

17. Discussion of expanding this figure to thirty million books while expending \$750 million has also been mentioned on the internet, but has not been confirmed by Google’s official website. See JONATHAN BAND, AMERICAN LIBRARY ASSOCIATION, OITP TECHNOLOGY POLICY BRIEF: THE GOOGLE LIBRARY PROJECT: THE COPYRIGHT DEBATE 10 (Jan. 2006), <http://ala.org/ala/washoff/oitp/googlepaprfnl.pdf>.

18. See Olsen, *supra* note 3. However, the capture methods used by Google produce extremely high-quality raw image files. Google Book Search, Information for Publishers and Authors About the Library Project, [http://print.google.com/googleprint/publisher\\_library.html](http://print.google.com/googleprint/publisher_library.html) (last visited Mar. 13, 2006) [hereinafter Information for Publishers].

19. Olsen, *supra* note 3, at 2.

20. *Id.*

21. *Id.*

22. *Id.*

23. See Information for Publishers, *supra* note 18 (answering the question: “How do I upgrade a Snippet View book into the Partner Program?”).

by Google's search engine, the distinction between copyrighted and public domain books will dictate the amount of information that will ultimately be available to a user as a search result.<sup>24</sup>

In response to a given query, Google Book Search will return a list of books that include the term within the books' text. If a book from the list is under copyright, a user will be presented with three "snippets" of text,<sup>25</sup> a count of the number of times the search term appears in the volume and links to online booksellers and information about the nearest local library that carries a print version of the book.<sup>26</sup> A user will be provided with only three snippets no matter how many times the search term comes up in the text.<sup>27</sup> Furthermore, snippets will not be provided for reference books like dictionaries, because such snippets could harm the market for the physical dictionary itself.<sup>28</sup> But if the work is in the public domain, a user will be able to browse the full text of the book.<sup>29</sup>

#### D. Google's Controversial Copyright Strategy: Opt-Out

Copyright still subsists in more than eighty percent of the materials in the Google 5 collection.<sup>30</sup> If Google were to request licenses for scanning and digitizing every print work from the work's copyright holder, not only might Google find it difficult to get a copyright holder's permission, but Google would likely also encounter problems with "orphan works,"<sup>31</sup> or

24. Information for Publishers, *supra* note 18.

25. A "snippet" of text consists of approximately three lines of content. *Id.*

26. *Id.* Although there will be no explicit identification of the library from which the book was scanned, public domain books, which are fully searchable, will display the respective library stamp and barcode imprinted inside the front and back of the books. UM Library/Google Digitization Partnership FAQ, Aug. 2005, at 6, <http://www.lib.umich.edu/staff/google/public/faq.pdf>.

27. See Band, *supra* note 17.

28. *Id.*

29. Information for Publishers, *supra* note 18.

30. Brian Lavoie et al., *Anatomy of Aggregate Collections: The Example of Google Print for Libraries*, D-LIB MAGAZINE (Sept. 2005), <http://www.dlib.org/dlib/september05/lavoie/09lavoie.html>. Google uses 1922 as a break-off date that demarcates those books that are in-copyright from those books that are out-of-copyright, or in the public domain. Information for Publishers, *supra* note 18 (answering the question: "How do you Determine if a Book is in the Public Domain and Therefore out of Copyright?"). This date is used for books written and published in the United States; outside the U.S. the break-off date may vary. *Id.*

31. Olive Huang, Note, *U.S. Copyright Office Orphan Works Inquiry: Finding Homes for the Orphans*, 21 BERKELEY TECH. L.J. 265 (2006); Tim Wu, *Leggo My Ego: Google Print and the other Culture War*, SLATE, Oct. 17, 2005, <http://www.slate.com/id/2128094>. Orphan works are any copyrighted works where the rights holder is hard to find. Because the cost of finding the owner is so high, creators cannot build on

those works whose authors are hard to find. In such a scenario, Google would fall extremely short of its goal to make “the full text of all of the world’s books searchable by anyone.”<sup>32</sup>

Understandably, Google has not requested permission for use of the books included in the project from any of the books’ copyright owners. Rather, Google has implemented an “opt-out” approach, in which copyright holders must notify Google if they do not want their work included in Google’s searchable library database.<sup>33</sup> This approach may seem contrary to copyright law since the one who copies or distributes a copyrighted work bears the burden of requesting permission from a copyright holder. However, when search engines index content, they do not formally request permission for such use despite the fact that such indexes copy entire webpages. Google maintains that, even absent permission, its Library Project is compliant with all copyright laws.<sup>34</sup>

Google believes such an approach to copyright should apply in the Google Library Project because trudging through millions of works and requesting permission title-by-title would be unwieldy and would generate prohibitive transaction costs. Moreover, the full-text copy kept in the Google database will never be accessible to users. Accordingly, Google is placing the burden on copyright holders to opt their works out of the Google Library search engine. This “opt-out” practice is anathema to some publishers and authors who point out that copyright in the print context is not an “opt-out” regime, and that the procedural rules of U.S. copyright law must be respected.<sup>35</sup> Author and publisher dissatisfaction with Google’s copyright approach has recently become the subject of two high-profile legal battles.<sup>36</sup>

---

orphan works, even when they would be willing to pay to use them. In many cases the works were abandoned because they no longer produced any income. In most cases, rights holders, once found, are delighted to have their work used. Save Orphan Works, <http://eldred.cc/> (last visited Mar. 13, 2006) (postings to Eldred blog).

32. Posting by Smith, *supra* note 14.

33. Forget Google Print Copyright Infringement: Search Engines Already Infringe, Posting of Danny Sullivan to SearchEngineWatch.com (May 25, 2005), <http://blog.searchenginewatch.com/blog/050525-093716>.

34. Google Print and the Authors Guild, Posting of Susan Wojcicki to Google Blog (Sept. 20, 2005), <http://googleblog.blogspot.com/2005/09/google-print-and-authors-guild.html>.

35. Under 17 U.S.C. § 106(1), the owner of a copyright has the exclusive right to reproduce the copyrighted work. 17 U.S.C. § 106(1) (2000).

36. On September 20, 2005, less than one year after the Google Library Project was announced, the Authors Guild, which represents more than 8,000 authors and is the largest society of published writers in the United States, filed a lawsuit, alleging that Google’s scanning and digitizing of library books constitutes “massive” copyright in-

## E. Google's Competitors and Their Library Projects

The uncertainty surrounding such copyright questions seems to have encouraged a wait-and-see approach by other technology companies who might engage in similar projects. Most large, legally savvy technology companies are hedging their bets and, while entering the rat race that has become the digital library, they are conservatively creating projects that clearly comply with current copyright law and utilize the traditional opt-in policy. In September 2005, Yahoo! and its partners announced their own "library project" which digitizes books, but only those in the public domain, except where the copyright holder has expressly given permission.<sup>37</sup> Meanwhile, Amazon announced its own version of a digital library project on November 3, 2005 with "Amazon Pages," a program that allows users to "buy-the-book" or "pay-per-page" online, from a selection of licensed works.<sup>38</sup>

The Google Library Project is different from the Yahoo! and Amazon projects in that it will provide access to books based on the great library collections of the United States. These library collections span public domain, orphan works, and copyrighted materials and are not limited to a particular group of publishers. Furthermore, Google has put its money where its mouth is—the cost of the Google Library Project is almost as enormous as its scope and supersedes the investment outlaid by its competitors' projects. Google has set aside \$150 million,<sup>39</sup> or about \$10 per

---

fringement. Elinor Mills, *Authors Guild Sues Google Over Library Project*, CNET-NEWS.COM, Sept. 20, 2005, [http://news.com.com/Authors+Guild+sues+Google+over+library+project/2100-1030\\_3-5875384.html](http://news.com.com/Authors+Guild+sues+Google+over+library+project/2100-1030_3-5875384.html). The Authors Guild demanded that Google pay damages for each infringement (up to \$150,000 per infringing copy) and asked the court for an injunction prohibiting the company from scanning copyrighted books without explicit permission. *See* Authors Guild v. Google, No. 05 CV 8136 (S.D.N.Y. Sept. 20, 2005). Nearly one month later, on October 19, 2005, the Association of American University Publishers on behalf of the McGraw-Hill Companies, Pearson Education, Penguin Group (USA), Simon & Schuster and John Wiley & Sons, sued Google for copyright infringement as well. *See* Association of American Publishers Sues Google over Library Digitization Plan, Posting of Gary Price to SearchEngineWatch.com (Oct. 19, 2005), <http://blog.searchenginewatch.com/blog/051019-115424>. The later suit seeks a declaratory judgment rather than an award for damages. *Id.*

37. Members of this project include the Internet Archive, Adobe Systems, the European Archives, the Biodiversity Heritage Library, the Smithsonian Institution Libraries, Hewlett Packard, Microsoft, Columbia University, and others. *See* Olsen, *supra* note 3.

38. *Amazon Takes Page from Apple*, *supra* note 3. Google is working on a model of pay-per-view charging as well. *Pulp Friction*, *supra* note 3.

39. Keith Regan, *Google Print Faces More Opposition*, E-COMMERCE TIMES, Aug. 30, 2005, <http://www.ecommercetimes.com/story/na7XBnDNb9BueY/Google-Print-Faces-More-Opposition.xhtml>.

book, for the initial digitization of the Google 5 collection.<sup>40</sup> Business models to recoup the investment are nascent and include paid advertisements at the bottom of a search page.<sup>41</sup>

Expending tens of millions of dollars to create this potentially valuable resource would be a worthy investment if the project were sure to be a success. However, the Google Library Project's potential for success is diluted by the uncertainty of copyright law's application to its project. In addition, there is no consensus on this project, as early commentators have posited an array of opinions. Some argue that the Library Project will make equal access to information a reality after a history in which knowledge has been maintained by the elite;<sup>42</sup> others suggest that the ability to perform Boolean searches of all the world's texts will revolutionize research and scholarship;<sup>43</sup> while others maintain that the social benefit is overstated and ultimately Google will merely recreate the traditional library card catalogue online.<sup>44</sup> Without consensus and clear precedent, the viability of the Google Library Project, thus, becomes a question of law for the courts.

## II. THE GOOGLE LIBRARY PROJECT IN A CHILD-SIZE COAT

In the context of current copyright law, the legality of Google's Library Project is questionable. Current lawsuits have alleged that the Library Project infringes on the copyright of every book it scans and digi-

40. However, the cost of the Google project, because of the company's technical capabilities and manpower, is significantly lower than the cost of similar library initiatives conducted by universities or smaller-scale technology companies. For example, if the University of Michigan were to proceed with its private digitization program that scans and digitizes nearly 5,000 volumes per year, it would take the library more than a thousand years and billions of dollars to digitize their entire collection. Google, on the other hand, forecasts that the University of Michigan digitization program will be complete in six years. UM Library/Google FAQ, *supra* note 12, at 2.

41. *Pulp Friction*, *supra* note 3. Google speculates that the project's financial value, based on advertising revenue generated by the project, will more than make up for the company's initial outlay. Burt Helm, *A New Page in Google's Books Fight*, BUSINESSWEEKONLINE, June 22, 2005, [http://www.businessweek.com/technology/content/jun2005/tc20050622\\_4076\\_tc119.htm](http://www.businessweek.com/technology/content/jun2005/tc20050622_4076_tc119.htm).

42. Mary Sue Coleman, *Creating a Global Online Library will Spread Knowledge in the Quickest Way to the Most People*, THE OREGONIAN, Nov. 13, 2005, at C1.

43. Matt Villano, *'Opening' a Digital Library*, CAMPUS TECH., Sept. 2005, available at <http://www.campus-technology.com/article.asp?id=11727>.

44. Heidi Benson, *A Man's Vision: World Library Online*, S.F. CHRON., Nov. 22, 2005, at A1, available at <http://www.sfgate.com/cgi-bin/article.cgi?file=/c/a/2005/11/22/MNGQ0FSCCT1.DTL>; see also Wu, *supra* note 31.

tizes without the permission of the author or copyright holder.<sup>45</sup> However, Google views the situation differently, retorting that the doctrine of fair use permits its scanning and digitizing activities, particularly when considering the enormous public good that will result from the successful completion of the project.<sup>46</sup> Three distinct activities of the Google Library Project raise copyright concerns under the current copyright regime.<sup>47</sup>

The first activity is Google's scanning and digitizing of full texts of copyrighted books from the Google 5 collection into Google's search database, effectively creating intermediate copies of millions of copyrighted works. The second activity arises in response to user queries. Google's search engine will present users with snippets, or a few sentences from the database's stored text, based on the user's search query.<sup>48</sup> The third activity involves Google's distribution of a copy of the digitized text to each of the five partner libraries. Although all three actions generate interesting questions for copyright law, this Note will focus, as do the pending lawsuits brought forth by the AAUP and the Authors Guild,<sup>49</sup> on the intermediate copy generated by digitization of the Google 5 collection as well as the snippets that a user will view when using the Google Book Search.

This Part will address the exclusive rights of copyright holders under the Copyright Act. Next, it will provide a background of the doctrine of fair use, and present relevant precedent. It will then apply the four-factor fair use test to Google's Library Project.

### A. Exclusive Rights of Copyright Holders

Owners of copyright in print books enjoy the exclusive right to make and distribute copies of that work, to "prepare derivative works based upon the copyrighted work," and to display and perform the work publicly.<sup>50</sup> In order to prove infringement, a copyright holder must demon-

---

45. See *Authors Guild v. Google*, No. 05CV8136 (S.D.N.Y. Sept. 20, 2005); Price, *supra* note 36.

46. Posting of Susan Wojcicki, Vice President, Google Product Management, to Google Blog (Sept. 20, 2005, 17:04 PDT), <http://googleblog.blogspot.com/2005/09/google-print-and-authors-guild.html>.

47. Band's article notes that Google's Library Project raises two concerns: one related to the scanning and digitizing and the other related to the snippets. See Jonathan Band, *Copyright Owners v. The Google Print Library Project*, E-COMMERCE LAW & POLICY, Aug. 2005, available at <http://www.policybandwidth.com/doc/googleprint-EntLaw.pdf>.

48. Band dismisses the snippets as *de minimis*, but notes that even if a court did not find the snippets to be *de minimis* they would then be subject to fair use analysis. *Id.*

49. See Mills, *supra* note 36; Price, *supra* note 36.

50. 17 U.S.C. § 106 (2000); 2 Melville B. Nimmer & David Nimmer, Nimmer on Copyright 8.02[d] (2005).

strate ownership of the copyright and establish that the alleged infringer copied original elements of the work.<sup>51</sup> The original work must be fixed in a tangible medium, but the copy need not be in the same medium as the original to constitute copyright infringement.<sup>52</sup>

While the Google 5 libraries may own the physical books that Google will scan and digitize, the authors and—in many instances—publishers control the copyrights of the books not yet in the public domain. Without obtaining a license from copyright holders, Google’s scanning and digitizing of books for the Google Library Project constitutes copying. However, the affirmative defense of fair use permits copying in certain circumstances, particularly those where value is being added to a copyrighted work. Ultimately, the legality of the copying that has and will occur by the Google Library Project rests on whether Google’s use constitutes a fair use.

## B. Fair Use

### 1. Background

The origins of the fair use defense to copyright infringement, currently codified in 17 U.S.C. § 107, date back to the early seventeenth century and the creation of the copyright.<sup>53</sup> Early courts recognized that certain instances of unauthorized reproduction of copyrighted material would not infringe the author’s rights.<sup>54</sup> A century later, Justice Story in a nineteenth century opinion laid out an approach for analyzing a question of fair use: “In short, we must . . . look to the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects of the original work.”<sup>55</sup> Today the Copyright Act includes the codification of a four-factor test for analyzing fair use, which largely reflects Justice Story’s approach.<sup>56</sup>

Despite its long history and codification, some argue that the fair use doctrine lacks predictability.<sup>57</sup> The confusing doctrine stumps judges, whose fair use decisions are often based on “intuitive reactions to

51. Feist Publ’ns v. Rural Tel. Serv. Co., 499 U.S. 340, 361 (1991).

52. See 17 U.S.C. §§ 101, 106 (2000).

53. See Act for the Encouragement of Learning, 1709, 8 Ann., c. 19 (Eng.).

54. See, e.g., Gyles v. Wilcox, (1740) 26 Eng. Rep. 489. See generally WILLIAM F. PATRY, THE FAIR USE PRIVILEGE IN COPYRIGHT LAW 6-17 (1985).

55. Folsom v. Marsh, 9 F. Cas. 342, 348 (C.C.D. Mass. 1841).

56. See 17 U.S.C. § 107 (2000)

57. See Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1107 (1990).

individual fact patterns,” and, as a result, “writers, historians, publishers and their legal advisers can only guess and pray as to how courts will resolve copyright disputes.”<sup>58</sup> Commentators assessing the validity of Google’s fair use defense disagree as to its legality.<sup>59</sup> Such is the legal quagmire in which Google and its Library Project are entangled.

## 2. *Relevant Precedent*

Two decisions, often quoted in commentator analysis, consider fair use in the context of digitization and provide some insight into the likelihood of success of Google’s fair use defense.

The Ninth Circuit considered fair use issues relating to search engine operation in *Kelly v. Arriba Soft Corp.*<sup>60</sup> In *Kelly*, a photographer sued a visual search engine for displaying thumbnail images of photographs originally posted on his website.<sup>61</sup> The Ninth Circuit found in favor of the search engine, holding that the search engine’s creation of thumbnails of the photographer’s copyrighted images, although used for commercial purposes, was a transformative, nonexploitative use and therefore fair.<sup>62</sup>

Another instructive example for Google’s project is the *MP3.com* case.<sup>63</sup> There, MP3.com purchased compact discs, created digital copies, and then stored the files in an online database. The company’s nonpaying subscribers were allowed to access any music that they could “prove” they owned or that they agreed to purchase.<sup>64</sup> When sued for infringement by a number of record companies, MP3.com maintained that the impact of its service would be positive since the service promoted purchase and ownership of the music.<sup>65</sup> The court rejected this argument on the grounds that the record companies had the right to grant or withhold a license to per-

58. *Id.*

59. See Elisabeth Hanratty, *Google Library: Beyond Fair Use?*, 2005 DUKE L. & TECH. REV. 10 (2005); Band, *supra* note 47; see also Posting of Michael Warnecke to OldFox Blog, Google’s Search of Library Shelves Splits Copyright Law Experts on Fair Use (Oct. 27, 2005), [http://oldfox.blogspot.com/2005\\_10\\_01\\_oldfox\\_archive.html](http://oldfox.blogspot.com/2005_10_01_oldfox_archive.html).

60. 336 F.3d 811 (9th Cir. 2003). Despite a substantial reliance on *Kelly* by many legal experts involved in the Google Library Project discussion, others question the long-term feasibility of *Kelly*: “It’s not at all clear that *Kelly* will withstand litigation in the Second Circuit.” Warnecke, *supra* note 59. If *Kelly* is in fact an out-of-the-mainstream Ninth Circuit decision, a more conservative judgment at the district court level could “hurt the viability of *Kelly*.” Warnecke, *supra* note 59.

61. *Kelly*, 336 F.3d at 816.

62. *Id.* at 822.

63. See *UMG Recordings, Inc. v. MP3.com, Inc.*, 92 F. Supp. 2d 349, 349 (S.D.N.Y. 2000).

64. *Id.* at 350.

65. *Id.* at 352.

form such a service.<sup>66</sup> The licensing market “directly derives” from the exclusive right granted to a copyright holder and the copyright holder has the right to “curb the development of such a derivative market by refusing to license a copyrighted work or by doing so only on terms the copyright owner finds acceptable.”<sup>67</sup>

The Google Book Search service provides services similar to those provided by Arriba Soft and MP3.com. The snippets presented to a Google Library user are comparable to the thumbnails in *Kelly*. A difference between the two activities, however, is that Google converts its print works into digital format and places them in an electronic database, while Arriba did not store a digital copy of the full-sized photographs. *MP3.com* is also distinguishable from the Library Project because Google is not using its digital copy for the purpose of supplanting the original use of the works as did MP3.com by allowing users to access music from their computers.<sup>68</sup> Neither of the cases is exactly on point, but both address fair use within the context of digital technology and are therefore instructive in analyzing Google’s fair use defense.

### 3. *Application of the Four-Factor Fair Use Test to Google’s Library Project*

The Copyright Act specifies four factors for courts to consider in analyzing a fair use defense: (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.<sup>69</sup> Other relevant factors, such as the “transformative” or creative nature of a use are also considered since “fair use is an ‘equitable rule of reason’ to be applied in light of the overall purposes of the Copyright Act.”<sup>70</sup>

---

66. *Id.* at 353.

67. *Id.* at 352.

68. *See infra* Section II.B.3.d.

69. 17 U.S.C. § 107 (2000).

70. *MP3.com*, 92 F. Supp. 2d at 350 (citing *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 448, 454 (1984)); *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 562-63 (1985); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 585 (1994).

a) The Purpose and Character of the Use

The first statutory factor listed in § 107 is “the purpose and character of the use.”<sup>71</sup> The transformative nature of a secondary use is also considered under this factor. It is likely that a court would find that this factor weighs in favor of fair use protection for Google’s Library Project.

i) Commercial Nature of the Use

The purpose and character of the use under the first factor considers “whether such use is of a commercial nature or is for nonprofit educational purposes.”<sup>72</sup> While Google created the Library Project for commercial gain, it is not attempting to profit from the sale or distribution of full-text copies of the books scanned and digitized into its database. Moreover, commercial use is no longer deemed by courts to be presumptively unfair.<sup>73</sup> The court in *MP3.com* held that more important than *MP3.com*’s commercial use was the fact that it exploited its secondary use of the copyrighted music by superseding the need for the original.<sup>74</sup> In the instant case, Google is not exploiting its use of the books by selling them online or making the print-copy available to users. Even though it will likely profit from advertising displayed alongside search results, Google’s primary objective is to make books accessible by making them easier to find. Search results provided by Google will serve to aid users in determining whether they are interested in pursuing the work further, i.e., “find[ing] the original at the library or purchas[ing] it after determining the work’s relevance to the user’s search.”<sup>75</sup> In this regard, the snippets are analogous to the thumbnails in *Kelly*, which provided users with just enough visual image for users to discern interest in the artist’s original photograph.<sup>76</sup> The Ninth Circuit held that such use was not “highly exploitative” of the artist’s work.<sup>77</sup>

---

71. 17 U.S.C. § 107(1).

72. *Id.*

73. *See Campbell*, 510 U.S. at 584.

74. *MP3.com*, 92 F. Supp. 2d at 351.

75. Hanratty, *supra* note 59.

76. *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818 (9th Cir. 2003).

77. *Id.* Thumbnails are arguably different from snippets in that snippets are exact quality reproductions of a work, and thumbnails are lower-quality reproductions of a work. However, snippets provide such minimal reproduction that they, like thumbnails, do not serve to replace full text/full resolution versions of the work. In this regard, a court may deem snippets to be de minimis and not infringing on a work’s copyright.

## ii) Transformative Nature of the Use

Under the first statutory factor, the Supreme Court has also examined the transformative nature of a secondary use and concluded that, “the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.”<sup>78</sup> Copied matter is transformed in the creation of “new information, new aesthetics, new insights and understandings.”<sup>79</sup>

A court would likely find the Google Library Project to be transformative. Although the search engine copies the works verbatim, digitizing and indexing books does more than shift printed material online. The activities also permit the conversion of content from the printed page to the internet, thereby creating new access through search features and new sales opportunities through exposure to book titles. A Google search leads users to relevant book titles and then guides them to a library, the publisher or an online bookseller. In short, Google’s copies serve a new purpose: the location and retrieval of information.

The Google Library Project is a technological middle ground between the utopia of free access to the world’s best libraries and the reality that authors and publishers must be paid or they will lose some incentive to write and publish books.<sup>80</sup> As courts de-emphasize the importance of commercial use and highlight the importance of the transformative nature of a use, Google’s case is strengthened. Accordingly, the first factor of a fair use inquiry would likely weigh in favor of Google.

## b) The Nature of the Copyrighted Work

Regarding the second factor, “the nature of the copyrighted work,”<sup>81</sup> a court would likely hold that this factor weighs against the Google Library Project. Under this factor, courts examine whether a copyrighted work is factual or creative, and creative works are given greater protection.<sup>82</sup> However, published works are protected less generously because copyright provides that it is the author’s right to first disseminate her work.<sup>83</sup>

Analyzing “the nature of the copyrighted work” is complicated in the instant case. Due to the fact that Google is scanning and digitizing mil-

---

78. *Campbell*, 510 U.S. at 579.

79. Leval, *supra* note 57, at 1111.

80. See Kevin Maney, *Critics Should Grasp Google Projects Before Blasting Them*, USATODAY.COM, Nov. 8, 2005, available at [http://www.usatoday.com/money/industries/technology/maney/2005-11-08-google\\_x.htm](http://www.usatoday.com/money/industries/technology/maney/2005-11-08-google_x.htm).

81. 17 U.S.C. § 107(2) (2000).

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

83. *Salinger v. Random House, Inc.*, 811 F.2d 90, 97 (2d Cir. 1987).

lions of books ranging from creative to fact-based, different types of works will likely be considered collectively. Creative works indexed by Google should be afforded a higher standard of protection than other more fact-based books. Nevertheless, the books, like the photographs in *Kelly*, will not receive heightened protection because they have either been published or made publicly available by the libraries that house them. The fact that the books may not be published on the internet is likely irrelevant. The Copyright Act defines “copies” as “material objects . . . in which a work is fixed by any method now known or later developed, and from which the work can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.”<sup>84</sup> Accordingly, the medium through which a work is transmitted is largely inconsequential, so long as the works are substantively the same. Given the creative nature of many works digitized by Google, it is likely that a court would view this factor as weighing against a finding of fair use.

c) The Amount and Substantiality of the Portion Used

The third statutory factor of § 107 is “the amount and substantiality of the portion used in relation to the copyrighted work as a whole.”<sup>85</sup> This factor seems to weigh against Google, since Google is copying the full text of millions of copyrighted books, and, in effect, making the entirety of each book available to a user in the form of a snippet. The court in *MP3.com* denied a claim of fair use under this factor on the same grounds—MP3.com copied the full text of CDs and presented the entirety of that text to its users. However, a court could also likely find, as did the Ninth Circuit in *Kelly*,<sup>86</sup> that this factor is neutral, weighing neither for nor against a finding of fair use because the secondary use necessitates use of the entire copyrighted work. Without full-text copying, Google would not be able to produce its robust database, and users would be unable to access snippets. Thus, the Google Library Project must copy the entire text of print works in order for its project to come to fruition. Accordingly, the court would balance the fact that, on the one hand, Google is copying the full text of print works, against the fact that Google *must* copy the full text of print works, and likely conclude that this factor does not weigh for or against a finding of fair use.

---

84. 17 U.S.C. § 101.

85. 17 U.S.C. § 107(3).

86. *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 821 (9th Cir. 2003).

#### d) The Effect of the Use on the Market

The fourth statutory fair use factor is “the effect of the use upon the potential market for or value of the copyrighted work,”<sup>87</sup> and represents the core of the publishers’ challenges to the Google Library Project. This factor considers whether unrestricted and widespread conduct of the sort engaged in by the purported fair user would result in a substantially adverse impact on the immediate or potential market for, or value of, publishers’ and authors’ copyrighted works.<sup>88</sup> Publishers argue that the Google Library Project will adversely impact the market for and value of books. Publishers claim that, accordingly, the Google Library Project challenges the utilitarian concept underlying copyright which promises creators the opportunity to realize rewards in order to encourage them to create. A secondary user that interferes excessively with an author’s incentives subverts the aims of copyright. Accordingly, the fourth factor disfavors a finding of fair use only when the immediate market is impaired because the copied material serves as a substitute for the consumer.<sup>89</sup> This factor also disfavors a finding of fair use when a potential market for a copyright holder is usurped by the secondary use.

#### i) The Immediate Market

Contrary to the publishers’ assertions, a court would likely find that this fourth factor weighs in favor of fair use. By leading users to the publishers or booksellers, Google—like Arriba Soft in *Kelly*—would actually guide users to legitimate sources of copyrighted material, likely driving sales of those titles.<sup>90</sup> Thus, the Google Library Project would not replace the need for physical print books, but would have a positive impact on the immediate market for physical books because users would buy more books, visit libraries, and use online bookseller sites.

---

87. 17 U.S.C. § 107(4). Prior to *Campbell*, the Supreme Court characterized the fourth factor as “the single most important element of fair use.” *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 566 (1985). Abandoning the idea that any factor enjoys primacy, *Campbell* instructs that “all [four factors] are to be explored, and the results weighed together, in light of the purposes of copyright.” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 (1994).

88. See 4 NIMMER, *supra* note 50, § 13.05 [A][4].

89. Leval, *supra* note 57, at 1125 (citing *Salinger v. Random House*, 650 F. Supp. 413, 425 (S.D.N.Y. 1986)).

90. For example, sales of Amazon.com’s searchable titles are 9% higher than others that are not searchable. Dan Richman, *New Amazon Search Feature Angers Authors*, SEATTLE POST-INTELLIGENCER, Oct. 31, 2003, available at [http://seattlepi.nwsourc.com/business/146262\\_amazon31.html](http://seattlepi.nwsourc.com/business/146262_amazon31.html).

The publishers' argument that the Google Library Project will replace the need for physical books is perplexing. The presence of the full-text copy on Google's internal database does not provide access to books. The database exists only so that it can be searched by the Google search engine in order to ultimately present snippets to users. While authors and publishers have argued that users could piece together snippets and eventually reproduce an entire work, that argument seems unrealistic, due to the difficulty and time-consuming nature of copying a work in such a manner. In addition, the Google Library Project creators have mitigated the possibility that users will reproduce works, through disabling the print, copy, and e-mail functions. Another counterargument that could be raised is that researchers may benefit from the snippets themselves, because they could glean valuable information from even the small amount of text presented. In such a case, in the absence of the Google Library search engine, a researcher would still likely not consult the actual books in the library, but rather rely on card catalogues or library computer databases to discover information similar to that which the Google Library would present.

It is plain that the Google Library Project does not replace the need for books; it merely indexes them. It does not serve the consumer as a substitute for books. If programs such as the Google Library Project can keep long-form written communication relevant, authors and book publishers will be better off.<sup>91</sup> The pages and card catalog information can only make people aware of books, not steal sales from authors and book publishers.<sup>92</sup> Historically, content owners, protective of their copyrighted works, often do not know what is good for their business. From John Philip Sousa opposing recorded music in the early 1900s,<sup>93</sup> to movie studios opposing the VCR nearly one century later,<sup>94</sup> content owners have opposed many new technologies that resulted in substantial profits in the long run.

---

91. See Maney, *supra* note 80.

92. *Id.*

93. John Philip Sousa warned in 1906 that:

[s]weeping across the country . . . comes now the mechanical device to sing for us a song or play for us a piano, in substitute for human skill, intelligence, and soul. . . . I foresee . . . injuries to music in its artistic manifestations by virtue—or rather by vice—of the multiplication of the various music-reproducing machines.

John Philip Sousa, *The Menace of Mechanical Music*, APPLETON'S MAG., Sept. 1906, at 278.

94. See *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984).

## ii) The Practice Becoming Widespread

While companies like Yahoo! and Amazon may prefer the wait-and-see approach at this point, if the Google-Publisher litigations are adjudicated in favor of fair use, similar projects would stampede the market.<sup>95</sup> The possibility of widespread use, though intimidating to publishers and copyright holders, would still not have a negative impact on immediate or potential markets. Widespread development of copyright-compliant digitization projects would increase potential exposure of titles that Google's collection does not include. Accordingly, such projects would promote the sale of more and different books, while encouraging library patronage by making users aware of the library location of the physical books.<sup>96</sup>

## iii) Usurping Potential Markets

Courts also address the possibility that a secondary use usurps a potential market for copyrighted works.<sup>97</sup> In the instant case, the publishers cannot claim that Google is usurping their potential market because, while there is a market for the sale of e-books or full-text books online, there is simply no potential market for the uses Google makes of the books at issue. Google is not selling books online, but rather creating an electronic card catalogue. The two activities are distinct. Unlike MP3.com, which usurped a potential market for record companies by selling and sharing full content online—a potential market for record distribution—Google's impact on potential markets for copyrighted works is negligible because publishers and copyright owners have little interest in developing an electronic card catalogue. The book publishing business came to computerization late and still seems to distrust anything digital. Many book publishers copy-edit 500-page books by hand on printouts.<sup>98</sup> Assuming, *arguendo*, that publishers were interested in such a potential market, it is uncertain that publishers would have the right to claim that the electronic card catalogue market is theirs to harness. None of copy-

---

95. Yahoo!, hoping to oust rival Google, has so far digitized mounds of material on its own. *Yahoo! to Digitize Library Contents*, WEB HOST INDUST. REV., Oct. 11, 2005, available at <http://www.thewhir.com/find/articlecentral/story.asp?recordid=1460&page=1>.

96. Perhaps an increase in library patronage would increase library budgets, which have been on the decline for decades, and also increase the acquisition of print books commensurately.

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

98. Maney, *supra* note 80.

right's exclusive rights suggest that publishers or authors should possess a monopoly over the indexing and searching of their works.<sup>99</sup>

While digitization projects like the Google Library Project could facilitate the growth of e-book markets, the Google Library Project does not attempt to compete directly with such markets. In fact, there seems to be a market for licensed content alongside the Library Project.<sup>100</sup> Accordingly, widespread indexing of copyrighted books would not harm the immediate or potential markets for print books, and this factor would likely weigh in favor of Google.

e) Overall Balancing

It is likely that the Google Library Project could be deemed a fair use, but detailed analysis of the four fair use factors fails to provide certainty regarding the likely success of Google's defense. Ultimately, Google's product offers more than a digital reproduction of massive library collections, by allowing extensive searching capabilities and access that were impossible in the print context. Additionally, the Google Book Search Project provides an out for parties who do not want to be involved with its "opt-out" policy, which permits any copyright owner or holder to have its content permanently suppressed from any search result.<sup>101</sup> Finally, there is a strong public interest in allowing this project to go forward.<sup>102</sup>

In *Sony*, a case involving hotly contested innovative technology, the court ruled in favor of increasing the rights of the public and technology developers rather than those of the copyright holders.<sup>103</sup> Ultimately, both sides won—the public got its VCR and the copyright holders cashed in on the popularity of the machine. Perhaps a similar outcome will result from the Google litigation. The inherent unpredictability of fair use determinations, however, suggests another possible resolution. Congress could resolve the question with legislation that re-examines and revises copyright law.

---

99. See 17 U.S.C. § 106 (2000).

100. Band, *supra* note 47, at 5 ("The existence of the Print Publisher Program, which involves licensing, demonstrates that the Library Project does not preclude lucrative licensing agreements.").

101. Wu, *supra* note 31.

102. See Tim O'Reilly, *Search and Rescue*, N.Y. TIMES, Sept. 28, 2005, at A27 ("Obscurity is a far greater threat to authors than copyright infringement, or even outright piracy . . . Google promises an alternative to the obscurity imposed on most books.").

103. See *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 417 (1984).

### III. LAWMAKERS: LETTING OUT THE HEM

Copyright law developed in response to technological change, specifically the development of the printing press, and as innovations have occurred, “Congress . . . has fashioned the new rules that new technology made necessary.”<sup>104</sup> Perhaps this is what is needed now as well. Google’s project should serve as an impetus to “clean up the copyright system.”<sup>105</sup> While revamping copyright law in Congress, like litigation, would also be a time-consuming undertaking, it would allow a permanent policy change so that future projects by private parties would not face the uncertain legal status that Google’s project faces.<sup>106</sup>

Legislators could provide initiatives like the Google Library Project with a little more breathing room by enacting legislation that would effectively create a narrow safe-harbor for digital library indexing projects. Congress should look to its own example of the Family Movie Act of 2005.<sup>107</sup> This Act permitted the makers of filtering software and components to manufacture and distribute their technologies without incurring liability. Congress could carve out a niche for indexing library collections as well. There is a pressing need for Congress to address a technological innovation that serves the public without harming copyright holders. Libraries exist to preserve society’s cultural artifacts and to provide access to them. If libraries are to continue to foster education and scholarship in this era of digital technology, it is essential for them to extend those functions into the digital world.<sup>108</sup> Congressional action would resolve the issue of intermediate copies in digitization projects, and allow socially beneficial projects to proceed.

#### A. Proposed Legislation: A Digital Library Safe Harbor

The Family Movie Act provides guidance in creating a safe harbor for digital indexing projects.<sup>109</sup> Rather than limiting the immunity of the Family Movie Act to specific companies or groups that were permitted to participate in creating and selling the filtering technology under the exemption, the Act defines the scope of the safe harbor by describing the precise

---

104. *Id.* at 431.

105. See Lawrence Lessig, *Let a Thousand Googles Bloom*, L.A. TIMES, Jan. 12, 2005, at B11.

106. See *id.*

107. Pub. L. No. 109-9, 119 Stat. 218 (2005).

108. Internet Archive, *supra* note 10.

109. See generally Alison R. Watkins, Note, *Surgical Safe Harbors: The Family Movie Act and the Future of Fair Use Legislation*, 21 BERKELEY TECH. L.J. 241 (2006).

action that is permitted.<sup>110</sup> Similarly, a digital library safe harbor could permit not only Google but other companies like it to participate. An ideal safe harbor would immunize:

- 1) the creation of an intermediate copy of full text digital content, so long as the copy is not distributed and reasonable steps are taken to secure the full text digital content; and
- 2) the use of internet search engine technology to access and search said intermediate copy, so long as only small amounts of text ultimately are presented to search users.

Such a safe harbor would immunize internet search engines from claims of copyright infringement for copying, indexing, and distributing or displaying to end users so long as reasonable measures are taken to secure the full-text digital content. Failing to take reasonable measures to secure the intermediate copy, or failure to present only limited content to users would disqualify a company from the safe harbor, rendering it liable for its negligence or resulting infringement.

## **B. Addressing Security Concerns**

To a greater extent than the technology at issue in the Family Movie Act, a digital library safe harbor generates serious concerns about security breaches of the full-text digital intermediate copy.<sup>111</sup> Google has announced that its digital copy will be kept in a safe archive utilizing advanced digital rights management technologies. In this archive, public domain works would be “lighted” and searchable by the public, whereas works under copyright would be kept in the “dark archive” until they become public domain.<sup>112</sup> Although the safe archive employs the latest advancements in anti-circumvention security, nothing digital is impervious to hackers who are intent on cracking security devices or strategies. To address such security issues, a hypothetical safe harbor could require that a certain percentage (e.g., 5%) of advertising revenues generated by the library projects be siphoned into an “insurance fund.”

First, such a fund would provide guaranteed financial recourse to a copyright holder whose copyrighted material was pirated or leaked. Sec-

---

110. *Id.* at 250-53.

111. *See Needless Fight Threatens Google's Online Library*, USA TODAY, Nov. 7, 2005, at 12A, available at [http://www.usatoday.com/news/opinion/editorials/2005-11-07-our-view\\_x.htm](http://www.usatoday.com/news/opinion/editorials/2005-11-07-our-view_x.htm); *see also Controversy Dogs Google's Book Search Library Project*, TAIPEI TIMES, Nov. 27, 2005, available at <http://www.asia-media.ucla.edu/article-us.asp?parentid=34337>.

112. *See* UM Library/Google FAQ, *supra* note 12.

ond, it would serve to insulate the search engine that took reasonable steps to secure its digital copy from liability, in case a hacker was able to penetrate it. Finally, such a requirement would impose a kind of “tax” on companies creating digital libraries, and thereby may dissuade smaller, less tech-savvy companies from participating in these projects, effectively discouraging the creation of unsecured intermediate copies. Such a fund has been integral in reaching legislative compromise in the past. Royalties were imposed on blank audio tapes, home recording devices, and other related products under the Audio Home Recording Act of 1992.<sup>113</sup>

### C. Legislation Versus Litigation

A narrow safe harbor would not address copyright’s failure to keep pace with technology in its entirety. A legislatively enacted safe harbor is, however, a better solution than litigation under the current regime for a number of reasons. First, litigation is both costly and time-consuming, and the outcome is uncertain. Second, a project is more likely to be enjoined if the court finds against fair use.<sup>114</sup> When a court rejects a fair use defense, it should deal with the issue of the appropriate remedy on its merits, granting or denying the injunction only for valid reasons, not simply as a mechanical reflex to a finding of infringement. Putting the brakes on digital library indexing this early in the game may stymie further innovation in the field.

Third, damages resulting from a court’s finding of massive copyright infringement could potentially bankrupt Google. Given the number of works at issue, the punitive effect of statutory damage awards could grow

---

113. 17 U.S.C. § 1008 (2000). The Audio Home Recording Act (AHRA) was a legislative compromise to deal with certain categories of digital audio copying. The Recording Industry Association of America (RIAA) was concerned that consumers could make perfect digital copies of music, thus destroying the market for audio recordings, and so it lobbied Congress to pass legislation. The AHRA required manufacturers of covered devices to (1) register with the Copyright Office; (2) pay a statutory royalty on each device and piece of media sold; and (3) implement a serial copyright management system which prevents all but first generation copies. In exchange for this, the manufacturers of the devices, which might have otherwise found themselves subject to liability for contributory copyright infringement, received a statutory immunity from suit. *See* RIAA Clarifies the Legality of the Home Audio Recording, <http://www.minidisc.org/ahra.html> (last visited Mar. 1, 2006).

114. *See* ALAN LATMAN, LATMAN’S THE COPYRIGHT LAW 278 n.105 (William F. Patry ed., 6th ed. 1986) (“Legal rhetoric has dulled thought on the injunction remedy. It is a vulnerable maxim that irreparable injury is ‘presumed’ in a case of copyright infringement.”).

excessive.<sup>115</sup> Moreover, if a court were to find that Google's use was not fair, it would likely also deem Google's behavior to have been willful, which, in turn, requires an award of maximum damages to plaintiffs.<sup>116</sup> Aggregated damages distort incentives to sue, and likely inform the publishers' decisions to bring suit against the Google Library Project. It is not that the publishers and copyright holders specifically object to Google's Project, as much as they object to being left out of their perceived share of the profits.<sup>117</sup> In such a lawsuit, the copyright holders would have undue recourse. By asserting their right to damages, authors and publishers could ultimately be handsomely compensated by Google, to the ultimate detriment of both Google and the public.

Legislation, in contrast to litigation, could define clear-cut rules and require that companies like Google abide by them.<sup>118</sup> A digital library safe harbor would solve the immediate question of how companies like Google, Yahoo!, Amazon, and others will proceed with their digital library projects today. However, such a fix will not solve a host of additional copyright issues that are perpetuated by the digital age and are certain to arise in the near future.

---

115. 17 U.S.C. § 504(b) grants the copyright owner his "actual damages suffered . . . and any profits of the infringer that are attributable to the infringement." 17 U.S.C. § 504(b) (West Supp. 2002). The copyright holder is permitted, however, to elect instead statutory damages of \$750 to \$30,000 per work infringed. *See id.* § 504(c)(1).

116. If the infringement was committed willfully, this statutory award may be increased to \$150,000. It may be reduced to \$200 if the infringer in certain narrow categories believed on reasonable grounds that his use was a fair use. *See id.*

117. Consumer book publishing is one of the most mature industries in media. *See Pulp Friction, supra* note 3. From 2003 to 2004, the number of books sold worldwide dropped by 44 million. Maney, *supra* note 80. Books are losing out to the internet, video games, DVDs, and podcasts, both in terms of the amount of time that people dedicate to them and the amount of money that people spend on them. *Pulp Friction, supra* note 3.

118. The likelihood that such legislative solutions will come to fruition is small. Publishers and copyright holders represent a powerful and institutional lobby on Capitol Hill, and former Congressional members are amongst the lobbyists fighting digitization for the publishers. Patricia Schroeder, President of the Association of American Publishers, one group that has brought a suit against Google for its library project is a near 20-year Congressional veteran. *See* [OpenSecrets.org, Former Members Turned Lobbyists](http://72.14.207.104/search?q=cache:GhpGfvQPX0gJ:www.opensecrets.org/pubs/lobby00/former.asp+Members+of+Congress+Publishers+Lobbyists&hl=en), <http://72.14.207.104/search?q=cache:GhpGfvQPX0gJ:www.opensecrets.org/pubs/lobby00/former.asp+Members+of+Congress+Publishers+Lobbyists&hl=en> (last visited Jan. 23, 2006). Such a powerful presence might overshadow any efforts of companies like Google, who just sent its first lobbyist to Washington D.C. in early October 2005. *Mr. Google Goes to Washington*, RED HERRING, Oct. 6, 2005, <http://www.redherring.com/Article.aspx?a=13899&hed=Mr.+Google+Goes+to+Washington>.

#### IV. THE GOOGLE TECHNICOLOR DREAMCOAT: RETHINKING CURRENT COPYRIGHT LAWS

Rethinking copyright law—an ideal solution for Google—would serve to remedy a broad host of issues that have cropped up as technology has outgrown copyright law. More sweeping legislative reform—in keeping with the theoretical underpinnings of copyright law—could address a variety of situations outside this particular scenario. This Part advocates a colorful alternative to § 106’s black letter law. By eliminating the exclusive right to reproduce as an organizing principle of copyright law, and putting in its place the exclusive right to distribute, the “copy” would be taken out of copyright.<sup>119</sup> Such restructuring would address the real commercial threat that unlicensed distribution of copyrighted works poses in the digital age.

##### A. The Purpose of Copyright

Any change in copyright law must maintain alignment with the purposes and goals of copyright protection. The Constitution grants Congress the power to enact laws that “promote the progress of Science and useful Arts.”<sup>120</sup> Promoting the creation and dissemination of knowledge furthers the public good. The copyright system, as Congress has styled it, aims to create incentives for authors to create and publishers to distribute new works. However, while these incentives are generally in the form of monetary remuneration, financial benefit derived by authors and publishers is not the end goal of copyright.<sup>121</sup> Copyright protects the exclusive rights of authors in order to allow them to recoup the costs associated with creativity and publishers to pay for distribution costs.<sup>122</sup> The mere occurrence of copying by users without distribution of those copies does not harm publishers and authors. Accordingly, distribution to the public is the necessary condition for harm to the publishers’ economic interests.

##### B. Reinvigorating Copyright by Focusing on Its Purpose

Before even the printing press, reproduction was a good predictor of intent to distribute.<sup>123</sup> Today, digital documents are easy to copy, and

---

119. See Ernest Miller & Joan Feigenbaum, *Taking the Copy Out of Copyright*, in SECURITY AND PRIVACY IN DIGITAL RIGHTS MANAGEMENT: ACM CCS-8 WORKSHOP DRM 2001, at 233, 238 (Tomas Sander ed., 2002), available at <http://www.cs.yale.edu/homes/jf/MF.pdf>.

120. U.S. CONST., art. I, § 8, cl. 8.

121. See Miller & Feigenbaum, *supra* note 119, at 234.

122. *Id.*

123. *Id.*

physical copies are easily digitized. In fact, in the computer world, copying is necessary in order to make any use of a work,<sup>124</sup> and allowing “normal use” of a work is a principle that copyright has traditionally supported. Reproduction in itself is no longer clearly indicative of intent to distribute. For example, people make copies of CDs that they have purchased as back-ups in case the original copy becomes scratched. This is distinct from making extra copies of purchased CDs to distribute to an entire group of friends, or worse yet, to sell to strangers on the subway. In the digital world, controlling copying is less important than controlling access to a work. Accordingly, copyright law should be rewritten to focus only on preventing distribution to the public, or to “a substantial number of persons outside of a normal circle of a family and its social acquaintances” borrowing from the language defining “to perform or display a work ‘publicly’” in the Copyright Act itself.<sup>125</sup> If the point of copyright is to prevent competition with those who have statutory rights, then legislation should protect that right (the right to distribute), and not promote a system that impedes “normal use” and technological advancement.

Replacing copyright with a distribution right would represent a significant change in intellectual property law and one that would not go uncontested. Content owners are possessive of their ownership interests in copyrights, and have made aggressive moves to protect them. The RIAA has considered a defense strategy which targets not just distributors, but also individual users who have copied MP3 files over the internet.<sup>126</sup> While the suggested alteration to copyright law seems dramatic, it would not undermine efforts of content owners to continue to seek out copyright infringers. For example, in the peer-to-peer context, most users distribute *and* copy. Targeting the distributors would also do away with a great majority of the users.

If Congress were to make such a change to the Copyright Act, Google would not face copyright infringement claims because its full-text digital copy would not be distributed to the public. Inaccessible to the public, Google’s intermediate full-text copy would not infringe any right of copyright. Beneficial library digitization projects could proceed without having to clear the substantial and outdated hurdles imposed by the current copyright regime.

---

124. See *Forget Google Print Copyright Infringement*, *supra* note 33.

125. 17 U.S.C. § 101 (2000).

126. See *RIAA’s Next Target Could Be You*, MEGAGAMES.COM, July 5, 2002, <http://www.megagames.com/news/html/pc/riaasnexttargetcouldbeyou.shtml>.

## V. CONCLUSION

Chances are that the copyright issues at stake in the Google Library Project will be adjudicated in the near future and such litigation will be dominated by two major questions battled out before a judge: Is this copyright infringement? Or, is this fair use? And, unfortunately, due to the unpredictability of the fair use doctrine, and a lack of precedent, there is no certainty about how a court will resolve the matter. What is certain is that a publishing house bringing suit against Google is not in the battle to uphold its constitutional right “to promote the Progress in Science and useful Arts,” but rather to obtain what it perceives to be its fair share of the Google Library Project’s profits. Although financial remuneration is *not* the end goal of copyright, it is the standard incentive model pursuant to Congress’s enactment of the Copyright Act. Accordingly, under current copyright law, incentives and progress go hand-in-hand. Publishers publish books to make money. But those books only promote progress if we read them. We can only read them if we can find them. The Google Library Project advances the public interest by making information globally accessible regardless of a user’s income, geographic location, and proximity to a library. In this way, it facilitates progress in science and the arts. The Project also simultaneously drives publishers’ incentives to create by increasing their profits based on increased exposure to book titles. Thus, the Google Library Project is consistent with copyright law and deserves legislative consideration.

