

2006

Copyright and Digitization Initiatives

Alicia Brillon

Follow this and additional works at: https://digitalcommons.law.uw.edu/law-lib_borgeson

Recommended Citation

Alicia Brillon, *Copyright and Digitization Initiatives* (2006), https://digitalcommons.law.uw.edu/law-lib_borgeson/7

This Article is brought to you for free and open access by the Law Librarianship Program at UW Law Digital Commons. It has been accepted for inclusion in Borgeson Paper Archive by an authorized administrator of UW Law Digital Commons. For more information, please contact lawref@uw.edu.

COPYRIGHT AND DIGITIZATION INITIATIVES

**By
Alicia Brillon**

**Current Issues in Law Librarianship, LIS 595
Spring Quarter 2006**

**Submitted to Professor Penny Hazelton
University of Washington School of Law
Seattle, Washington**

May 8, 2006

Contents

Copyright Issues	1
Public Domain	1
Fair Use.....	3
Five Current Digitization Projects	4
Yale's Avalon Project.....	4
Carnegie Mellon's Posner Memorial Collection	5
Personnel Issues	6
Determining Copyright Status	6
Foreign Works and Copyright	7
Determining Copyright Ownership.....	7
The Results.....	7
Project Costs	7
Conclusion	8
Google Print.....	8
Google Print Publisher.....	8
Google Book Search	9
Opt-Out Procedure	10
The Positive Aspects.....	11
Evaluation	11
Google's Stance On Scanning	13
Google's Stance On Snippets.....	14
Yahoo! and the Open Content Alliance	15
Yahoo! vs. Google	16
Amazon.com Search Inside the Book	16
Publishers Reaction.....	17
Lessons To Be Learned.....	17
Newest Digitization Initiatives	18
Should Libraries Undertake Digitization Efforts?	19
Yes, With A Commercial Venture.....	19
Issues To Address When Pushing the Legal Boundaries.....	20
Yes, But Do It On Their Own.....	20
Yes. With Other Law Libraries.....	21
 Flowchart for Determining When.....	 Appendix A
U.S. Copyright In Fixed Works Expire	

There are many organizations that have recently embarked on projects to digitize written materials, either for preservation or to make the information available via the Internet, or both. Some of these organizations are libraries or similar non-profit institutions, while others are commercial ventures. Each has adopted a different strategy for determining what types of materials will be digitized, with varying degrees of success and short-comings. Five such initiatives will be discussed: Yale's Avalon project, Carnegie Mellon's Posner Memorial Collection project, Google Print, Yahoo! and the Open Content Alliance, and Amazon.com's Search Inside the Book project. The various difficulties encountered and an analysis regarding whether law libraries should embark on such efforts will be covered.

Copyright Issues

To understand the issues surrounding these current digitization projects, a review of copyright law is needed to set the stage. Under United States copyright law, the owner of a copyrighted work has the exclusive right to control how their work is used, including the reproduction, creation of derivative works, distribution, and performance of the work.¹ However, this right to control is not absolute. It is subject to an exception for uses that meet a four-pronged fair use test.²

Public Domain

Fair use is not the only way that permits someone to use a work without permission of course. Works that are in the public domain are not subject to copyright provisions and can be used freely.³ There are a number of ways that a work can become owned by the public, ie. "public domain," as opposed to being owned by the original author or another party.

First, an author can purposefully place a work into the public domain. In this case, the author will specifically state within the work itself that no copyright is being claimed and that any one can make any use of the work. A statement such as "I dedicate this work to the public domain" would be appropriate. This is not common however.⁴

-
1. Title 17 of the United States Code contains the Copyright law.
 2. 17 U.S.C. § 107 (2000).
 3. LORNA M. HUGHES, *DIGITIZING COLLECTIONS: STRATEGIC ISSUES FOR THE INFORMATION MANAGER 60* (Marilyn Deegan and Simon Tanner eds., Facet Publishing 2004).
 4. Stanford University, *Copyright and Fair Use*, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter8/8-a.html#3 (last visited May 3, 2006).

Second, some works are not ever protected by copyright and are in the public domain upon creation.⁵ Included in this category are works created by federal government employees (in their capacity as an employee) such as Presidential speeches, government reports, statutes and court decisions.⁶

Third and most commonly, a copyright can simply expire, causing the work to fall into the public domain.⁷ All works published in the United States before January 1, 1923 are currently in the public domain due to expired copyrights. These works can be used in any way without having to obtain permission.⁸

Lastly, a copyright holder can fail to have observed a technicality regarding the copyright, causing the copyright to lapse and allowing the work to fall into the public domain. This could have happened in two ways.

First, for works published in the United States after 1922 but prior to 1964, the copyright holder was required to file a renewal with the Copyright Office during the 28th year after the work was first published. If this renewal was not timely filed, the copyright would be lost.⁹ United States copyright law was simplified in 1992 when Congress enacted a law that provided for automatic renewal of works copyrighted between January 1, 1964 and December 31, 1977.¹⁰

Second, if a work was first published in the United States prior to March 1, 1989, it must have included a copyright notice. If this notice was not included, or was incorrect as far as form is concerned, copyright protection was lost. Note that if the work was first published between January 1, 1978 and March 1, 1989, the copyright holder could follow prescribed actions to cure the defect and preserve the copyright.¹¹

-
5. MELISSA SMITH LEVINE, HANDBOOK FOR DIGITAL PROJECTS: A MANAGEMENT TOOL FOR PRESERVATION AND ACCESS 70 (Maxine K. Sitts ed., Northeast Document Conservation Center 2000).
 6. Stanford University, *Copyright and Fair Use*, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter8/8-a.html#3 (last visited May 3, 2006).
 7. LORNA M. HUGHES, DIGITIZING COLLECTIONS: STRATEGIC ISSUES FOR THE INFORMATION MANAGER 60 (Marilyn Deegan and Simon Tanner eds., Facet Publishing 2004).
 8. Mary Minow, *Library Digitization Projects and Copyright*, http://www.llrx.com/features/_and_Fair_Use_digitization2.htm#95 (last visited May 2, 2006).
 9. Stanford University, *Copyright and Fair Use*, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter8/8-a.html#3 (last visited May 3, 2006).
 10. Copyright Renewal Act of 1992, Pub. L. No. 102-307, 106 Stat. 264. (amending chapter 3, title 17 of the *United States Code*, by providing for automatic renewal of copyright for works copyrighted between January 1, 1964, and December 31, 1977), enacted June 26, 1992.
 11. Stanford University, *Copyright and Fair Use*, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter8/8-a.html#3 (last visited May 3, 2006).

While the above may seem fairly straight-forward, trying to determine if a copyright has expired or a work is otherwise in the public domain is actually quite complicated, as the flowchart in Appendix A clearly shows.¹²

Fortunately, going forward, determining copyright status will be less complicated. Works which were created after January 1, 1978 are now protected for the life of the author plus 70 years (up from life plus 50 years), and no renewal terms are provided. The foregoing is the result of Congress' enacting the Sonny Bono Copyright Term Extension Act (CTEA) in 1998.¹³ This act provides that instead of being protected for a term of the life of the author plus 50 years, works created after January 1, 1978 are now protected for a term of the life of the author plus 70 years. However, this means that no works first published after 1978 will fall into the public domain until 2019.¹⁴

Fair Use

Some of the digitization efforts currently being undertaken have taken the stance that their actions are covered by the fair use exception to the Copyright Law. Google's Book Search project¹⁵ is the main proponent of this stance and that project will be discussed at length shortly. But first a brief summary of fair use is in order.

Fair use as an exception to the Copyright Law is provided for under 17 U.S.C. § 107 (2000). A four factor test as laid out by that section is employed to determine whether a particular use is permissible. The four factors are:

- (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.

12. Bromberg & Sunstein LLP, <http://www.bromsun.com/practices/copyright-portfolio-development/flowchart.htm> (last visited May 2, 2006).

13. Sonny Bono Copyright Term Extension Act, title I of Pub. L. No. 105-298, 112 Stat. 2827 (amending chapter 3, title 17, *United States Code*, to extend the term of copyright protection for most works to life plus 70 years), enacted October 27, 1998.

14. Mary Minow, *Library Digitization Projects and Copyright*, <http://www.llrx.com/features/digitization2.htm#95> (last visited May 2, 2006).

15. Google's Book Search project is discussed at page 8 of this paper.

A court will balance the above factors, none of which will determine whether a particular use is fair or not, and evaluate them under the specific facts of each individual case. If the balance weighs in favor of the party claiming fair use, it is defense to a claim of copyright infringement.¹⁶

Fair use can be found in a number of settings, with the most common being in either an educational or library setting. However, one case that is of particular interest for digitization projects is Kelly v. Arriba Soft, a case involving a commercial entity.¹⁷

The Kelly case involved a database search engine that reproduced thumbnail copies of images in its search results. Kelly, a commercial photographer, maintained that the thumbnail versions of his works displayed by the search engine were a violation of his copyrights.

However, the court found that the thumbnail images served an "entirely different function" than the original images which were high quality commercial photographs. The thumbnails served a public benefit through the search engine and even though the entity was commercial in nature their activities were deemed to fall within the realm of fair use.

Five Current Digitization Projects

The copyright and fair use discussions above highlight the intricacies of the issues involved when determining whether a work is covered by copyright, and if so, how long that protection remains in force. The impact of copyright law on digitization projects becomes clear when various projects are examined more closely. Following is an analysis of five such projects, each of which has varying approaches to dealing with the copyright issue.

1. Yale's Avalon Project

Yale's Avalon Project is an example of a straight-forward project that digitizes out-of-copyright material only. The Avalon Project contains information and documents from the fields of law, history, economics, politics, diplomacy and government for both the United States and foreign countries.¹⁸ Collections

16. MELISSA SMITH LEVINE, HANDBOOK FOR DIGITAL PROJECTS: A MANAGEMENT TOOL FOR PRESERVATION AND ACCESS 72 (Maxine K. Sitts ed., Northeast Document Conservation Center 2000).

17. Kelly v. Arriba Soft, 336 F.3d 811 (9th Cir. 2003).

18. Yale University Avalon Project Home Page, <http://www.yale.edu/lawweb/avalon/avalon.htm> (last visited May 2, 2006).

comprise a wide range of topics from “Nazi-Soviet Relations 1939 – 1941” to “United States Statutes Concerning Slavery.”

According to the project’s statement of purpose, it endeavors to not only provide the text of the applicable primary documents via the World Wide Web, but also to add value by linking to supporting documents which are referred to in the original text.¹⁹ The Avalon Project also includes links to full bibliographic records for the source documents included.

While the Avalon Project is not controversial as far as copyright concerns, it does state that it will likely contain controversial documents. “Their inclusion does not indicate endorsement of their contents nor sympathy with the ideology, doctrines, or means employed by their authors. They are included for the sake of completeness and balance and because in many cases they are by our definition a supporting document.”²⁰

2. Carnegie Mellon’s Posner Memorial Collection

In a paper detailing Carnegie Mellon’s efforts to digitize the Posner Memorial Collection, Denise Troll Covey outlined a number of challenges encountered.²¹ She then explained how Carnegie dealt with those challenges. This section summarizes the main points of Covey’s findings.

In 2001, Carnegie Mellon was given funding by Henry and Helen Posner to digitize the 1,106 works collected by Henry’s father from 1924 through 1973, known as the Posner Memorial Collection. This collection consists of fine and rare books on the topics of history, art and literature, as well as some related archival material. The funding would allow Carnegie Mellon to purchase a scanner specifically designed to work with fine and rare materials as well as to pay the scanner operator. Once digitized, the works would be made freely available via the Internet.

Given the date span during which the works in the Posner Memorial Collection were collected, it is not surprising that many of the works were covered by copyright. However, it is very interesting to note that Carnegie Mellon didn’t even question the fact that permission would need to be obtained prior to scanning

19. Yale University Avalon Project, <http://www.yale.edu/lawweb/avalon/help/helpdesk.htm#Purpose> (last visited May 2, 2006).

20. *Id.*

21. DENISE TROLL COVEY, ACQUIRING COPYRIGHT PERMISSION TO DIGITIZE AND PROVIDE OPEN ACCESS TO BOOKS, 21-38 (Council on Library and Information Resources 2005).

these works: “We knew that the collection contained some copyrighted titles and therefore that the project entailed acquiring copyright permission.”

Personnel Issues

With the copyright permission issue in mind, Carnegie set out to digitize the Posner Memorial Collection. One staff member was assigned to work on the project; however this person did not have a great amount of time to devote to it. Just over one year after beginning the project, Carnegie found that only 75 letters requesting permission to digitize copyrighted works had been sent out, and only one third of the publishers contacted had actually responded. It was determined that at the current rate of progress it would take four and one half years to complete the permission aspect of the project. Carnegie decided that additional personnel must be obtained in order to more efficiently complete the project.

Carnegie hired a part-time person in May of 2003 who would be dedicated to the permission work, and then extended her to full-time status in September of that year. Having a dedicated person moved the project along, and most of the permissions work was completed by November 2003. Still, there were lingering communications with various publishers through 2004.

Determining Copyright Status

Besides obtaining permissions, another aspect of Carnegie’s digitization project that made it time-consuming and difficult was simply trying to determine whether a given title was covered by copyright or not. (Keep in mind the flowchart cited earlier in this article.) If published between 1923 and 1963, when copyrights had to be formally renewed, the only way to find out if this had been done was to consult the Copyright Office for a title search. Carnegie decided to conduct a test, asking the Copyright Office to do title searches for seven titles. The Copyright Office charged \$150 upfront, and informed Carnegie that it would take four to six weeks for results to be obtained. In the end, the Copyright Office found only one of the seven titles.

Carnegie determined that it would cost from \$6,000 to \$8,000 to have the Copyright Office conduct title searches for all of the relevant materials in the Posner Memorial Collection, with no guarantee that results would be successfully obtained. Carnegie then decided that a better way to spend their time and resources would be to only research the copyright renewal records if the status of a particular work was in question.

Foreign Works and Copyright

Another aspect of the digitization project which caused Carnegie difficulties was trying to determine the copyright status of foreign works. Due to the patchwork of ways in which works are protected in various countries, it was exceedingly difficult to determine whether a particular work was protected by copyright or not. After initially working with Carnegie's legal counsel, it was decided that the foreign copyright issue was too complex and was slowing down the project in general. Carnegie then decided to simply assume that all of the foreign works were protected by copyright and obtain permission accordingly.

Determining Copyright Ownership

After going through the process of determining if a work was protected by copyright or not, the next obstacle for Carnegie was to determine who actually owned a given copyright. Intertwined with this issue was the problem of actually finding the copyright owner. The publisher or author listed on a work's title page may or may not be the current copyright owner. Publishers go out of business, copyrights are abandoned or assigned, and there is no one single source to consult for either ownership or location information. A lot of detective work was therefore involved in identifying and then tracking down the applicable copyright owner for each particular work. In those cases where Carnegie could not locate the copyright owner, it was assumed that permission was denied and the work would not be included in the digitization initiative.

The Results

At the end of 2004, Carnegie was still trying to locate nearly a third of the publishers for copyrighted works. Contact was attempted via various methods, including by letters, emails and telephone calls. Once contacted, 75% of the publishers granted permission for Carnegie to include their work in the digitization project. Some publishers explicitly did not grant permission. Others were deemed by Carnegie itself to have not given permission. This determination was made after either three attempts had been made to contact the publisher without a response, or if Carnegie was unable to locate the publisher at all. Carnegie decided to err on the side of caution and not assume that any publisher would want to be included in their digitization project.

Project Costs

Carnegie kept track of all of the costs involved with the digitization project. These consisted of personnel in the form of part-time and then full-time workers, as well as costs for paper, envelopes, postage and long-distance telephone fees. After factoring in all costs, Carnegie determined that each copyrighted work for which permission was granted cost \$78. This is a conservative number, however,

as Carnegie did not take into account additional costs such as the use of in-house counsel, Internet connectivity, database creation or the intermittent time devoted to the project by library personnel prior to May 2003.

Conclusion

In the end, Carnegie Mellon did complete their project and the full text of 622 titles in the Posner Memorial Collection is available online.²² However, this particular digitization project underscores that while it is possible to include copyrighted works in such an initiative, determining copyright status, identifying the copyright owner and locating that owner are all difficult and time-consuming steps. Having dedicated personnel to work on the project is an important factor, as is having a realistic view that some works may not be able to be included in the project in the end.

3. Google Print

Google Print was announced in December 2004. The project itself consists of two parts: first is Google Print Publisher, which is promoted as a way for publishers to get their works out in front of the buying public and increase sales.²³

The second part is Google Book Search (also know as Google Print Library or Google Books Library Project). This is promoted as “an enhanced card catalog of the worlds books.”²⁴ While Google Print Publisher has been welcomed by many publishers, as will be discussed below, Google Book Search has received much harsher scrutiny due to Google’s questionable interpretation of copyright law.

Google Print Publisher

Through Google Print Publisher, publishers can sign up to submit their books for inclusion in Google's search index.²⁵ It is important to note that under this side of the Google Print project, the publisher is given the choice of whether or not to participate. If a publisher chooses to be included, the full text of their works will

22. Carnegie Mellon Posner Memorial Collection Home Page,

<http://posner.library.cmu.edu/Posner/> (last visited May 3, 2006).

23. Google Inc., *Partner Program - An Online Book Marketing and Sales Program*, <http://books.google.com/googlebooks/publisher.html> (last visited May 3, 2006).

24. Google Inc., *Google Books Library Project - An Enhanced Card Catalog of the World's Books*, <http://books.google.com/googlebooks/library.html> (last visited May 2, 2006).

25. Google Inc., *Partner Program - An Online Book Marketing and Sales Program*, <http://books.google.com/googlebooks/publisher.html> (last visited May 3, 2006).

be scanned and indexed into Google's system.²⁶ Publishers will share in the revenue generated from the contextual ads that Google pairs with applicable search results,²⁷ and users can view the page on which their keyword search term is located, as well as two pages both forward and backward in the book from the page on which their search term appears.²⁸

Displayed with the search results are links to online retailers selling the book, including a link to the publisher itself.²⁹ Users are not able to print or download the text nor print or copy images as that functionality has been disabled.³⁰

Google Book Search

The response to Google Print Publisher from publishers and authors has been positive. However, Google Book Search is much more controversial. For this project, Google collaborated with five major libraries: the University of Michigan, Harvard, Stanford, Oxford, and the New York Public Library.³¹ These libraries are giving Google access to the content contained within their collections and in return, Google will scan the full text of the content at their expense (roughly ten cents per page) and provide a digital copy to the library that supplied the book.³² Google will keep a second copy for inclusion in their own database.³³ Note that one party is left out and will not receive a digital copy – the copyright holder, in many cases an existing and identifiable author or publisher.³⁴

26. Google Inc., *Your Content Is Protected*, http://books.google.com/services/print_tour/print4.html (last visited May 3, 2006).

27. Google Inc., *Earn New Revenue with Contextual Ads*, http://books.google.com/services/print_print_tour5.html (last visited May 3, 2006).

28. Google Inc., *Your Content Is Protected*, http://books.google.com/services/print_tour/print4.html (last visited May 3, 2006).

29. Google Inc., *Drive More Book Sales*, http://books.google.com/services/print_tour/print3.html (last visited May 3, 2006).

30. Google Inc., *Your Content Is Protected*, http://books.google.com/services/print_tour/print4.html (last visited May 3, 2006).

31. Google Inc. Press Release, *Google Checks Out Library Books*, http://www.google.com/press/pressrel/print_library.html (last visited May 3, 2006).

32. Jessica Dye, *Scanning the Stacks: the Digital Rights Issues Behind Book Digitization Projects*, ECONTENT, Jan.-Feb. 2006, at 32.

33. Id.

34. Barbara Quint, *About that Flaw in Google Print*, INFORMATION TODAY, Oct. 2005, at 7.

Opt-Out Procedure

In 2005, Google began scanning portions of these collections, ignoring the copyright status of the items. After various authors and publishers expressed their concern with this plan, Google did implement a mechanism for publishers to opt-out of participating in Google Book Search. The author/publisher can tell Google not to include a particular work in the project by supplying the applicable book's ISBN, and Google has stated that it will honor that request.³⁵

However, it appears that some requests have not been honored. Some authors say that their requests to be excluded have not been successful as the content in question is still online.³⁶ (This does not surprise me. Having worked at internet companies in the past, I know first hand how difficult it can be to permanently remove any particular piece of content from a database.)

The above opt-out system may seem a bit backwards to those who are familiar with copyright law, and its convention of putting the onus on the would-be copier to proactively gain permission from the copyright holder before making a copy of a work.³⁷ However, from Google's perspective, this is the way the online world works, or at least how it has worked in the past. An opt-out procedure is the method Google provides for the websites it indexes – anyone who does not want Google to include their website in Google's index can simply include an exclusion header or robots.txt file in their code.³⁸

Note how drastically Google's approach differs from the assumptions made by Carnegie Mellon with regards to the Posner Memorial Collection just a few years earlier, namely: 1) that permission must be obtained if a work is protected by copyright, and 2) never assume that a publisher would want to be included in Carnegie's digitization project absent explicit permission having been obtained.

The Positive Aspects

Despite the difficulties the Google Book Search project has encountered, many people were very enthusiastic when it went live, and indeed there are actually a number of positive aspects to the Google Book Search project.

35 Jonathan Band, *The Google Print Library Project: A Copyright Analysis*, <http://www.policybandwidth.com/doc/googleprint.pdf> (last visited May 3, 2006).

36 Matthew Hicks, *University Presses Slam Google Print*, EWEEK, May 24, 2005, <http://www.eweek.com/article2/0,1759,1820047,00.asp> (last visited May 4, 2006).

37 Stephanie Hodge, *Google Battles Over Books*, DAILY BRUIN, October 28, 2005.

38 Jonathan Band, *The Google Print Library Project: A Copyright Analysis*, <http://www.policybandwidth.com/doc/googleprint.pdf> (last visited May 3, 2006).

For instance, Google plans to include millions of out of print titles in their online database.³⁹ Many of these titles are currently available only from libraries, so access to these works would be greatly increased.⁴⁰ Google will also provide options to the book searching consumer as well - when a user searches and finds a book they like, they can click on any number of links to online sellers of that book on the results page. If the applicable publisher has given Google permission to show sample pages in response to a user query, the publisher's link will be the first link on the list.⁴¹ (If the publisher has not given permission, their link will not be included at all.) Clicking on a link will take the user directly to the detail page of the applicable online seller so that the book can be purchased immediately.⁴²

Google has also stated that a "Find this in a library" link will be provided on the results page. When a user clicks on that link, they will be taken to the OCLC WorldCat database. The user then enters their zip code in a search box and a list of nearby libraries that have the item in their collection is returned.⁴³ However, as of May 2006, the "Find this in a library" link was only present on the Google Book Search results pages for public domain works. There is speculation that publishers feel that the library link would decrease sales, so Google is not rushing to implement the link for copyrighted materials.⁴⁴

Evaluation

The above features certainly sound positive, and indeed they are. However, there are other aspects to the Google Book Search project that are causing more controversy. The main issue is Google's policy of scanning and including books in their project that are under copyright, but for which permission has not been affirmatively received from the author or publisher to include the work in the Google Book Search project.

Many authors and publishers feel that Google's actions are an infringement on

39. Richard A. Leiter, *Dodging and Weaving Through the Online Libraries: Focus On Google Print*, LEGAL INFORMATION ALERT, Sept. 2005, at 7.

40. Id.

41. Google Inc., *Drive More Book Sales*, http://books.google.com/services/print_tour/print3.html (last visited May 4, 2006).

42. Google Inc., *Google Book Search Help Center*, http://books.google.com/support/partner/bin/answer.py?answer=18625&topic=322&hl=en_US&gsessionid=x0U877oeJ5o (last visited May 4, 2006).

43. Google Inc., *Frequently Asked Questions*, <http://books.google.com/googlebooks/help.html#3> (last visited May 4, 2006).

44. Andrew Albanese and Norman Oder, *Google Gains with Librarians: Newsletter Sign-ups Boom; Book Sales Launch; Where's Library Link*, LIBRARY JOURNAL, Apr. 15, 2006, at 6.

their copyrights,⁴⁵ and not surprisingly, they have made their feelings known. The Association of American University Presses which consists of 125 scholarly publishers, first expressed their concerns in a letter to Google in May of 2005.⁴⁶ In response, Google did stop scanning copyrighted material temporarily and it was at this time during the summer of 2005 that they instituted the opt-out policy.⁴⁷

But apparently the opportunity to opt-out was not enough to placate the authors and publishers, because in September 2005 the 8,000 members of the Authors Guild filed a class action lawsuit against Google in federal court in Manhattan.⁴⁸ The Association of American Publishers followed with their own lawsuit filed in the same court one month later.⁴⁹ So far, none of the five libraries who have provided content to Google have been named in a lawsuit.⁵⁰

Both actions against Google cite two basic copyright infringement issues: 1) the initial act of scanning the materials, and 2) the reproduction of sections of the scanned books following a user's search.⁵¹ The later are referred to as "snippets" in Google's parlance.⁵²

Authors feel that Google's act of scanning and digitizing their works is a violation of their right as copyright owners to control who makes a copy of their works.⁵³ No permission means no copies.

-
45. Ben Charny, *Google Print Pressures Libraries*, eWeek, Sept. 23, 2005 pNA.
 46. Matthew Hicks, *University Presses Slam Google Print*, eWeek, May 24, 2005, <http://www.eweek.com/article2/0,1759,1820047,00.asp> (last visited May 4, 2006)
 47. Jessica Dye, *Scanning the Stacks: the Digital Rights Issues Behind Book Digitization Projects.*, ECONTENT, Jan.-Feb. 2006, at 32.
 48. The Authors Guild, *Authors Guild Sues Google, Citing "Massive Copyright Infringement,"* Sept. 20, 2005 http://www.authorsguild.org/news/sues_google_citing.htm (last visited May 5, 2006).
 49. Association of American Publishers, *Publishers Sue Google Over Plans To Digitize Books*, Oct. 19, 2005, <http://publishers.org/press/releases.cfm? PressReleaseArticleID=292> (last visited May 5, 2006).
 50. Ben Charny, *Google Print Pressures Libraries*, EWEK, Sept. 23, 2005.
 51. Jonathan Band, *The Google Print Library Project: A Copyright Analysis*, <http://www.policybandwidth.com/doc/googleprint.pdf> (last visited May 3, 2006).
 52. Google Inc., *"Authors: Common Questions,"* http://books.google.com/googlebooks/author_fa.html#4 (last visited May 3, 2006).
 53. No author, *"Publishers Sue Google Over Plans to Digitize Copyrighted Books: Google Print Library Violates Publishers' and Authors' Rights,"* US NEWSWIRE, Oct. 19, 2005.

Google's Stance On Scanning

Google on the other hand feels that their scanning of library materials is permitted. Google states that "The 'fair use' provisions of U.S. copyright law (USC 17 107) describe the conditions under which someone may copy a work without the copyright holder's permission..." They go on to explain that Google has "...carefully designed Google Book Search to make sure our use of books is fair and fully consistent with the law."⁵⁴

Additionally, Google equates the act of scanning a book to that of indexing a webpage. In order to index a webpage, says Google, you need to make a copy of it. Similarly, in order to index a book, you need to make a copy of it.⁵⁵

The problem with that argument in is that every web publisher knows that by posting content on the Internet, it is going to be indexed by search engines and will be freely available for viewing by anyone who finds it (password protected websites aside). It is commonly accepted that there is an implied license to index freely accessible websites.⁵⁶ But the same cannot be said about a book whose content has never been placed online by the author or publisher, which is the case with the majority of books. This is where Google's analysis breaks down.⁵⁷

Google will likely try to compare the Book Search project to the image database at issue in *Kelly v. Arriba Soft*, arguing that their scanning of the book material is transformative, making it a fair use. Google's argument would stress that the purpose of scanning is not to reproduce the original content of the book, but to transform the content and make it part of an index.⁵⁸

However, another case dealing with the creation of an online index of content occurred in 2000, when the Recording Industry Association of America

54. Google Inc., *Google Book Search: News & Views*, http://books.google.com/googlebooks/newsviews/facts_fiction.html (last visited May 5, 2006).

55. Google Inc., *Information for Publishers and Authors About the Library Project*, http://books.google.com/googlebooks/publisher_library.html (last visited May 5, 2006).

56. Jonathan Band, *The Google Print Library Project: A Copyright Analysis*, <http://www.policybandwidth.com/doc/googleprint.pdf> (last visited May 3, 2006).

57. Raymond Nimmer, *Google Lawsuit Begins; Fair Use*, <http://www.ipinfoblog.com/archives/intellectual-property-33-google-lawsuit-begins-fair-use.html> (last visited May 5, 2006).

58. Neil J. Rosini and Michael I. Rudell, *Google Print: Suits by Authors and Publishers Against Google Raise Fair Use Questions*, <http://www.fwrv.com/news/article.cfm?id=100645> (last visited May 5, 2006).

Issues To Address When Pushing the Legal Boundaries

In the later scenario, it would be wise for the library to negotiate adequate contractual protections with the commercial entity. For instance, provisions which provide for defense and/or indemnification by the commercial entity should a copyright claim be filed against the library would be wise. Additionally, a provision which allows the library to pull out of the project with either little or no notice period, and for any or no reason, would be advantageous.

Similarly, if a library decides to partner with a commercial venture, it should be prepared for possible negative reactions, either by the press or by author's trade groups. While so far there has not been great criticism leveled directly at the libraries participating in the Google initiative, it may just be a matter of time before they are implicated. And while it is doubtful that the Author's Guild or AAP would take joy in bringing a lawsuit against a library, at some point this may occur if necessary to fully protect what they see as their infringed rights.

Lastly, as has been the case with the University of Michigan in the Google project, a public university library should be aware that their contract with a commercial entity may become public via an FOIA request. The contract should be drafted and negotiated with this possibility in mind.

Yes, But Do It On Their Own

In the case where a library is not as concerned with making materials available to the entire world, but only wishes to preserve unique or rare materials in a more stable format or make particular content available to their immediate patrons, undertaking a digitization project on their own may be the best course of action.

In this scenario, the sum of the content will likely be fairly manageable, although as Carnegie's Posner Project demonstrates even smaller digitization projects can be very time-consuming. The simplest sort of digitization project for an individual library would be one that only involves out of copyright content, such as that in the Yale Avalon project. Staff time and equipment would be the primary costs. If a library wishes to include copyrighted material, they again would need to determine which side of the fence they fall on regarding whether scanning and digitizing this material is permissible absent the copyright holder's permission.

In either case, a library would be wise to study past digitization initiatives in order to become familiar with the issues involved, and to set realistic goals for both the total costs that will be incurred as well as how long the project will take.

Yes, With Other Law Libraries

Finally, work by the Legal Information Preservation Alliance (LIPA) provides a backdrop to the issue of digitization. Whatever format preservation may take, that valuable materials are preserved and made available for future generations is of the utmost importance. While LIPA has not undertaken any digitization initiatives to date, their work toward preserving legal information originally published in either print or electronic format can provide a guide to any law library that would like to ensure that unique or rare parts of its collection be accessible long after the original physical record embodying the material has been lost.

