

---

---

# The Googlization of Everything and the Future of Copyright

Siva Vaidhyanathan\*

*Based on an explicit rejection of the “revolutionary” claims that supporters have put forth in favor of full-text searching of centuries of books sitting in American university libraries, this Article argues that the legal challenges to Google’s plans to digitize millions of copyrighted books from university libraries are likely to succeed, thus hampering other, potentially more valuable public indexes and digitization processes. In addition, the assertions that Google and its supporters have made about the legality and potential benefits of the Google Book Search service are counterproductive to the maintenance of a healthy and open information ecosystem. The Google case is the most potentially disruptive copyright battle since the invention of sound recording technology. It strikes at the very heart of the copyright system and reveals that we tend to rely on the rickety structure of fair use to support too many essential public values. Google’s Library Project threatens to unravel everything that is good and*

---

\* Associate Professor of Culture and Communication, New York University. I must thank the following people for engaging me in deep and extended conversation on this topic in the preparation for this draft: Oren Bracha, Tony Reese, Cory Doctorow, Fred von Lohmann, Michael Madison, Lawrence Lessig, Laura Gasaway, Georgia Harper, Niko Pfund, Eric Zinner, Carol Mandel, Howard Besser, Laura Quilter, Jack Bernard, Farhad Manjoo, Alexander Macgillivray, Steve Maikowski, Carlos Ovalle, Liz Losh, Michael Zimmer, Aaron Swartz, Donna Wentworth, Peter Jaszi, Robert Boynton, Lawrence Weschler, Melissa Henriksen, Sam Howard-Spink, Richard Rogers, Jonathan Band, Robert Kasunic, Patricia Schroeder, Alan Adler, Sam Trosow, Catherine Collins, Chris Lydon, James Neal, Jessica Litman, Susan Douglas, and Helen Nissenbaum. Versions and portions of this work appeared as talks delivered to the 2005 Annual Meeting of the American Library Association, the University of Texas School of Law, a meeting of the Washington, D.C. Copyright Society, the Department of Culture and Communication at New York University, the Sweetwater Writing Center at the University of Michigan, Emory University Libraries, and the Swiss Technical University in Zurich. An early version of these arguments appeared in Siva Vaidhyanathan, *A Risky Gamble with Google*, CHRON. HIGHER EDUC. (Wash. D.C.), Dec. 2, 2005.

*stable about the copyright system. It injects more uncertainty and panic into a system that is already in disequilibrium.*

## TABLE OF CONTENTS

INTRODUCTION .....	1209
I. THE BASICS OF GOOGLE BOOK SEARCH.....	1215
A. <i>Privacy</i> .....	1219
B. <i>Privatization</i> .....	1220
C. <i>Property</i> .....	1221
II. THE ARGUMENTS FOR GOOGLE.....	1222
III. PRAGMATISM OR FUNDAMENTALISM? .....	1225
IV. A LAPSE INTO LEGAL REALISM .....	1226
V. WHY QUALITATIVE FACTORS MATTER.....	1227
VI. THE TRUMP CARD: THE LIBRARY COPY .....	1229
CONCLUSION.....	1230

## INTRODUCTION

In May 2006, *Wired* magazine contributor Kevin Kelly published in *The New York Times Magazine* his predictive account of flux and change within the book-publishing world. In that article Kelly outlined what he claimed “will” (not “might” or “could” — “will”) happen to the book business and the practices of writing and reading under a new regime fostered by Google’s plot to scan in millions of books and offer searchable texts to Internet users.

“So what happens when all the books in the world become a single liquid fabric of interconnected words and ideas?” Kelly asked.

First, works on the margins of popularity will find a small audience larger than the near-zero audience they usually have now. . . . Second, the universal library will deepen our grasp of history, as every original document in the course of civilization is scanned and cross-linked. Third, the universal library of all books will cultivate a new sense of authority.”<sup>1</sup>

Kelly sees the linkages of text to text, book to book, page to page, and passage to passage as the answer to the knowledge gaps that have made certain people winners and others losers. Kelly wrote:

If you can truly incorporate all texts — past and present, multilingual — on a particular subject then you can have a clearer sense of what we as a civilization, a species, do know and don’t know. The white spaces of our collective ignorance are highlighted, while the golden peaks of our knowledge are drawn with completeness. This degree of authority is only rarely achieved in scholarship today, but it will become routine.”<sup>2</sup>

Such heady, passionate predictions of technological revolution have become so common, so accepted in our techno-fundamentalist culture, that even when John Updike criticized Kelly’s vision in an essay published a month later in *The New York Times Book Review*, he did not doubt that Kelly’s universal digital library would someday come to pass. Updike just lamented it, musing about how wonderful his old bookstore haunts were for him and everyone else who strolled the streets of New York, Oxford, or Boston in the 1950s.<sup>3</sup> Updike’s

---

<sup>1</sup> Kevin Kelly, *Scan This Book!*, N.Y. TIMES, May 14, 2006, § 6 (Magazine), at 42.

<sup>2</sup> *Id.*

<sup>3</sup> See John Updike, *The End of Authorship*, N.Y. TIMES, June 25, 2006, § 7 (Sunday

elitist comments only served to bolster the democratic credentials of Kelly and others who have been asserting that Google's plan to scan millions of books would spread knowledge to those not as lucky as Updike.

As it turns out, the whole move toward universal knowledge is not so easy. Kelly's predictions depend, of course, on one part of the system Kelly slights in his article: the copyright system. Copyright is not Kelly's friend. He mentions it as a mere nuisance waking him from his dream of a universal library. But to acknowledge that a lawyer-built system might trump an engineer-built system would have run counter to Kelly's sermon. In fact, the current American copyright system will most certainly kill Google's plan to scan in the entire collections of the University of Michigan Library and more than a dozen other major American research university libraries. And that is not necessarily a bad thing.

Copyright in recent years has certainly become too strong for its own good. It protects more content and outlaws more acts than ever before. It stifles individual creativity and hampers the discovery of and sharing of culture and knowledge.<sup>4</sup> But the Google Library Project sits so far beyond the scope of traditional copyright that it threatens the very foundation of the law. Google is exploiting the instability of the copyright system in a digital age. It hopes to rest a huge, ambitious, potentially revolutionary project on the most rickety, least understood, most provincial, most contested perch among the few remaining public interest provisions of American copyright: fair use.

---

Book Review), at 27.

<sup>4</sup> See generally DAVID BOLLIER, *SILENT THEFT: THE PRIVATE PLUNDER OF OUR COMMON WEALTH* (2002); BENJAMIN KAPLAN, *AN UNHURRIED VIEW OF COPYRIGHT, REPUBLISHED (AND WITH CONTRIBUTIONS FROM FRIENDS)* (2005); LAWRENCE LESSIG, *CODE AND OTHER LAWS OF CYBERSPACE* (1999); LAWRENCE LESSIG, *FREE CULTURE: THE NATURE AND FUTURE OF CREATIVITY* (2004); LAWRENCE LESSIG, *THE FUTURE OF IDEAS: THE FATE OF THE COMMONS IN A CONNECTED WORLD* (1st ed. 2001); JESSICA LITMAN, *DIGITAL COPYRIGHT: PROTECTING INTELLECTUAL PROPERTY ON THE INTERNET* (2001); KEMBREW MCLEOD, *FREEDOM OF EXPRESSION: OVERZEALOUS COPYRIGHT BOZOS AND OTHER ENEMIES OF CREATIVITY* (1st ed. 2005); KEMBREW MCLEOD, *OWNING CULTURE: AUTHORSHIP, OWNERSHIP, AND INTELLECTUAL PROPERTY LAW* (2001); THE COMMODIFICATION OF INFORMATION (Niva Elkin-Koren & Neil Weinstock Netanel eds., 2002); SIVA VAIDHYANATHAN, *COPYRIGHTS AND COPYWRONGS: THE RISE OF INTELLECTUAL PROPERTY AND HOW IT THREATENS CREATIVITY* (2001); SIVA VAIDHYANATHAN, *THE ANARCHIST IN THE LIBRARY: HOW THE CLASH BETWEEN FREEDOM AND CONTROL IS HACKING THE REAL WORLD AND CRASHING THE SYSTEM* (2004) [hereinafter VAIDHYANATHAN, *THE ANARCHIST*]; Siva Vaidhyanathan, *Copyright as Cudgel*, *CHRON. HIGHER EDUC.* (Wash., D.C.), Aug. 2, 2002, at 7.

The peer-to-peer “crisis” was supposed to be the greatest threat to date to the historically successful copyright system and the industries that depend on it. As it turned out, the massive distribution of billions of copyrighted music files via the Internet did not destroy the commercial music industry. If anything, downloading strengthened the industry. More importantly, the behavior of more than seventy million people who offered and received copyrighted files without payment did not undermine the foundations of copyright. The system continues to work. Songwriters write. Producers produce. Distributors distribute. Lawyers sue. Downloaders download. We learned three essential truths from the downloading debate: a shared file is not a lost sale, there is a significant difference between a crisis and a moral panic, and culture is not zero-sum.<sup>5</sup>

In an amicus brief I wrote on behalf of media studies professors in what was supposed to be the landmark showdown over peer-to-peer file sharing and the future of technology, *Metro-Goldwyn-Mayer Studios, Inc. v. Grokster Ltd.*,<sup>6</sup> I argued that there was no functional distinction between the peer-to-peer interface Grokster and the popular search engine Google.<sup>7</sup> They are both search engines that facilitate discovery, access, and infringement of others’ copyrighted works. They both “free ride” on others’ copyrighted works. They both provide a service to the public for no direct remuneration through subscription, yet they are both commercial entities that benefit from increased traffic and the data gathered from their users. So if you want to find Grokster liable for inducing infringement, you have to consider that Google’s web search service is liable as well.<sup>8</sup>

Of course, there is one big difference: Grokster did not actually do any copying. Google does. For years Google has been making cache copies of web pages it indexes. Its search function cannot operate without a cache index on which to rely. In two cases, courts ruled that this practice of generating cache copies for the purpose of

---

<sup>5</sup> VAIDHYANATHAN, THE ANARCHIST, *supra* note 4, at 43-43, 48-50.

<sup>6</sup> 380 F.3d 1154, 1158 (9th Cir. 2004).

<sup>7</sup> Brief for Media Studies Professors as Amici Curiae in Support of Respondents at 4, 10, *Metro-Goldwyn-Mayer Studios, Inc. v. Grokster Ltd.*, 380 F.3d 1154 (9th Cir. 2004) (No. 04-480).

<sup>8</sup> Lawrence Lessig makes a similar point in one of his earliest Weblog posts on the Google Library Project, in which he pleads that if the Library Project were ruled to be massive infringement it might endanger all of Google’s enterprises. This is essentially my premise as well, although I derive a much different and conservative conclusion from it. See Posting of Lawrence Lessig to Lessig Blog, <http://www.lessig.org/blog/archives/003140.shtml> (Sept. 22, 2005, 5:37 EST).

enabling search engines is non-infringing.<sup>9</sup> The Ninth Circuit Court of Appeals's ruling in *Kelly v. Arriba Soft Corp.* established a strong measure of confidence for search engines and other web enterprises.<sup>10</sup> Yet the principle that caching web pages to enable search engines is non-infringing is far from settled, as copyright holders have taken aim at the Ninth Circuit's ruling in *Kelly*.<sup>11</sup> Still, wholesale copying for caches continues. It seems certain that we could not navigate the World Wide Web effectively if Google and other search engines could not freely copy and cache others' copyrighted material on the web.

Now, through its Library Project, an element of the Google Book Search service, Google plans on copying much more. Google is reaching beyond the Internet and into the real world. It hopes to impose the copyright norms of the digital world onto the analog world. Publishers, accustomed to the norms of the real world and skittish about those of the web world, are panicking and suing.<sup>12</sup> Google has put at risk not only the precedent established in *Kelly*. It has gambled the value of the company. It has, in the words of Michael Madison, "bet the Internet" on this case. If Google loses, its (and everyone's) ability to copy and cache material without a license would be in jeopardy. Google and other search engines make the Internet make some sense, allowing us to navigate in a rough and imperfect

---

<sup>9</sup> See *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 817-22 (9th Cir. 2003); *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1124 (D. Nev. 2006); see also *Google Free to Cache: Court*, RED HERRING (Belmont, Cal.), Jan. 26, 2006.

<sup>10</sup> See generally JOHN BATTELLE, *THE SEARCH: HOW GOOGLE AND ITS RIVALS REWROTE THE RULES OF BUSINESS AND TRANSFORMED OUR CULTURE* (2005) (explaining web search technology and its potential application to range of content); Khoi D. Dang, Note, *Kelly v. Arriba Soft Corp.: Copyright Limitations on Technological Innovation on the Internet*, 18 SANTA CLARA COMPUTER & HIGH TECH. L.J. 389 (2002) (explaining effects of ruling that thumbnail images are transformative and that effective search engines require presumed license to cache).

<sup>11</sup> Alan Adler, counsel for the American Association of Publishers, has told me that he considers *Kelly* to be a bad decision that he and other content industry representatives would like to see overturned. In early 2006, a district court in California ruled that caching images and producing thumbnail images to facilitate searching should not always be considered fair use. See *Perfect 10 v. Google, Inc.*, 416 F. Supp. 2d 828, 851 (C.D. Cal. 2006).

<sup>12</sup> See, e.g., *Author's Guild v. Google Inc.*, No. 05 CV 8136 (S.D.N.Y. filed Sept. 20, 2005); *McGraw-Hill Cos. Inc. v. Google Inc.*, No. 05 CV 8881 (S.D.N.Y. filed Oct. 19, 2005). Considerations of how norms influence copyright regulation and practices are essential to a full understanding of the relationship among cultural communities, industries, markets, and regulatory systems. See Ann Bartow, *Electrifying Copyright Norms and Making Cyberspace More Like a Book*, 48 VILL. L. REV. 13, 17 (2003).

way through a rapidly expanding set of documents. So there is much at stake.<sup>13</sup>

In this Article I will introduce some of the issues at stake in the Google Library Project. I want to emphasize that much of what I am about to say is speculative. I have tried to be as pragmatic and undogmatic as possible. I am no enemy of Google. I am a fan and a customer. I think it is one of the coolest companies to come along in my lifetime. I just think libraries are much cooler. And I worry that libraries might be hurting their standing by hanging out with the geeks.

This is about more than handicapping legal battles. There is much at stake in Google's ambitions. Google plays a peculiar and powerful role in our information culture. It is a ubiquitous brand, used as a noun and a verb everywhere from adolescent conversations to scripts for *Sex and the City*. Its initial public offering in 2004 generated \$1.67 billion in cash.<sup>14</sup> Its stock price soared in value immediately thereafter, only to erode by nearly thirty percent from its peak in the early months of 2006.<sup>15</sup> Its revenue has more than doubled to \$3 billion per year since the offering.<sup>16</sup> Because the company refuses to release quarterly income projections, investors have engaged in rumor mongering about the company, making it a constant subject of conversation in the financial press and the bar rooms near Wall Street. The core service of Google.com — its web search engine — handles just more than fifty percent of the web search business in the United States.<sup>17</sup> That is just ahead of Yahoo's share.<sup>18</sup> Google is much stronger overseas.<sup>19</sup> Yet Microsoft is gaining on both of them.<sup>20</sup>

To preserve its status as the elite, venerated, and fast-moving technology company of the future, Google must do two things. It must continue to convince the world that it is the anti-Microsoft. And it must find more things to index and expose to the world. Microsoft controls most of the desktops in the world. It also controls an

---

<sup>13</sup> Madison considers this a good bet. I do not. See Michael J. Madison, Madisonian.net, *Google Print II* (Oct. 20, 2005), <http://madisonian.net/archives/2005/10/20/google-print-ii/>.

<sup>14</sup> See Charles H. Ferguson, *What's Next for Google*, Tech. Rev., Jan. 2005, at 6, available at <http://www.technologyreview.com/Infotech/14065/>.

<sup>15</sup> *Id.*

<sup>16</sup> *Id.*

<sup>17</sup> *Id.*

<sup>18</sup> *Id.*

<sup>19</sup> *Id.*

<sup>20</sup> *Id.*

increasing number of operating systems for mobile data devices. So many of the world's files are potentially indexible and searchable by Microsoft itself. And Microsoft has ways of locking other firms out of essential services that concern the desktop environment. Microsoft also still dominates the browser "market." Thus, it can leverage its ubiquity and proximity to muscle in on the web search business. In addition, the chief advantage Google has had in the web search area, its clean and effective search algorithm, PageRank, is no longer the only such method for effective web searches. Other firms, including Microsoft and Apple, can now make good search engines. *Sex in the City's* Carrie Bradshaw might be a Google customer now, but there is no reason to believe she will be next year.<sup>21</sup>

Google has to protect its brand by being seen as the good guy on the block. And so far it has. The damage Google has done to the world is minimal and centers largely on the slippage of grammatical standards, encouraging more people to use its brand as a verb. Google got big by keeping ads small. It carefully avoided pinching our marketing-saturated nervous systems and offered illusions of objectivity, precision, comprehensiveness, and democracy.<sup>22</sup> After all, we are led to believe, Google search results are determined by peer review by us not by an editorial team of geeks.<sup>23</sup> So far, this method has worked wonderfully. Google is the hero of word-of-mouth marketing lore. There is a huge, annoying billboard in my neighborhood trumpeting Yahoo's influence in the world. Yet every day when I turn on my computer, I see the comforting white background and cheerful cartoon graphic imagery of Google instead. Google guides me through the open web, the space that Microsoft does not yet control. Just as clearly, Google must get bigger. It must go new places and send its spiders crawling through un-indexed corners of human knowledge. Google's mission statement includes the rather optimistic and humanistic phrase, "to organize the world's information and make it universally accessible and useful."<sup>24</sup> But Google cofounder, Sergey Brin, once offered a more ominous description of what Google might become: "The perfect search engine would be like the mind of God."<sup>25</sup>

---

<sup>21</sup> *Id.*

<sup>22</sup> See DAVID A. VISE & MARK MALSEED, *THE GOOGLE STORY* 96 (2005).

<sup>23</sup> See Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, 16 *INFO. SOC'Y* 169, 171-72 (2000).

<sup>24</sup> Google.com, Corporate Information: Company Overview, <http://www.google.com/corporate> (last visited Feb. 22, 2007).

<sup>25</sup> Ferguson, *supra* note 20, at 2.

## I. THE BASICS OF GOOGLE BOOK SEARCH

Since 2003, Google has been negotiating with and securing digital rights from commercial and academic publishers to support what used to be called “Google Print.” Thousands of images of book pages sit in Google’s servers under this authorized program. The terms of access to the images depend on the particular wishes of the publishers. Some offer nearly full-text access. Others offer only small portions of the books. In general, users may view only a few pages of a book at a time, and they may not print or download the images. The margins of the Book Search pages contain links to sources where one may purchase the books, as well as publication information and links to the publishers’ sites.

In December 2004, Google announced its plans to digitize millions of bound paper books from five major English-language libraries. Then, in August 2006, the University of California system announced its partnership with Google to scan its books.<sup>26</sup> Of these six libraries, only the University of Michigan and the University of California system will allow Google to scan their complete collections.<sup>27</sup> The libraries’ initial contributions are as follows:

- (1) Harvard University libraries: 40,000 public-domain books during the pilot phase of the project, which might be extended. The library has more than 15 million volumes.
- (2) Stanford University libraries: Hundreds of thousands of public-domain books, but officials eventually may allow Google to scan the entire collection of 7.6 million books.
- (3) University of Michigan at Ann Arbor: All 7.8 million books in the collection, even those under copyright.
- (4) University of Oxford: All books published before 1900. The library holds a total of 6.5 million books in its collection.

---

<sup>26</sup> Press Release, Jennifer Colvin, University of California Office of the President, UC Libraries Partner with Google to Digitize Books (Aug. 9, 2006), <http://www.universityofcalifornia.edu/news/2006/aug09.html>; see also Scott Carlson, *U. of California Will Provide Up to 3,000 Books a Day for Google to Scan*, CHRON. HIGHER EDUC. (Wash., D.C.), Sept. 8, 2006, at A32.

<sup>27</sup> See Scott Carlson & Jeffrey R. Young, *Google Will Digitize and Search Millions of Books from 5 Top Research Libraries*, CHRON. HIGHER EDUC. (Wash., D.C.), Jan. 7, 2005, at A37.

- (5) The New York Public Library: From 10,000 to 100,000 public-domain volumes as part of a pilot project. The library holds 20 million volumes.<sup>28</sup>
- (6) The University of California system: More than 2.5 million books at a scanning rate of 3,000 volumes per day. The books will come from more than 100 libraries on the 10 University of California campuses.

In total, Google plans to add more than 17 million library volumes to its electronic index at an estimated cost of \$10 per book, or \$170 million. As payment for access to the books, Google will provide the libraries with an electronic copy of the works they contributed to the project.<sup>29</sup>

Under the unauthorized library project — which is distinct from the “partner” project authorized by publishers — search results and the user experience depend on the copyright status of the book. For works published before 1923 (and thus safely in the public domain in the United States), the user will have access to the entire text. For works published since 1923 (and thus potentially under copyright protection), the user will see the bibliographic information as well as a few text “snippets” around the search term. Google claims that viewing the displayed results of copyrighted works is comparable to the “experience of flipping through a book in a bookstore.”<sup>30</sup> Google disables the user’s print, save, cut, and copy functions on the text display pages so that the user is limited to reading the information on the screen. As with the authorized “partner” content, Google provides links to “buy this book” from numerous vendors as well as targeted advertisements that will depend on the nature of the book and, conceivably, what Google assumes to be the interests of the searcher.

The initial reactions from publishers were unwarranted. They expressed concerns that the Google project would threaten book sales,

---

<sup>28</sup> See Michael Gorman & John P. Wilkin, *One College Librarian Worries About ‘Atomizing’ Books*, CHRON. HIGHER EDUC. (Wash., D.C.), June 3, 2005, at A25.

<sup>29</sup> See Cooperative Agreement § 2.5, Google-Univ. Mich., available at <http://www.lib.umich.edu/mdpum-google-cooperative-agreement.pdf> (last visited Feb. 22, 2007); see also Elisabeth Hanratty, *Google Library: Beyond Fair Use?*, 2005 DUKE L. & TECH. REV. 0010, ¶ 1 (2005), <http://www.law.duke.edu/journals/dltr/articles/PDF/2005DLTR0010.pdf>.

<sup>30</sup> Press Release, Google Inc., Google Checks Out Library Books (Dec. 14, 2004) [http://www.google.com/press/pressrel/print\\_library.html](http://www.google.com/press/pressrel/print_library.html); see also Google, Google Books, Library Project, <http://books.google.com/googlebooks/library.html> (last visited Feb. 22, 2007).

risk hacking, and the subsequent “Napsterization” of text.<sup>31</sup> As the conversation matured and the debate distilled certain matters, it became clear that Google’s Library Project offers no risk to publishers’ core markets and projects. If anything, the project could serve as a marketing boon (that is, if the searches actually generate quality results that help users identify the right books for the task at hand).<sup>32</sup> Since then, it has become clear that publishers now are most offended by the prospect of a wealthy corporation “free-riding” on their content to offer a commercial and potentially lucrative service without any regard to compensation or quality control. The publishers would like a piece of the revenue and some say in the manner of display and search results.<sup>33</sup> Copyright rarely has been used as leverage to govern ancillary markets for goods that enhance the value or utility of the copyrighted works. There have been some examples in recent years as the power and scope of copyright grew beyond any rational justification. But courts have recently looked upon such efforts with suspicion.<sup>34</sup> As author and activist Cory Doctorow has pointed out, booksellers have never insisted on extracting licensing revenues from bookcase makers, bookmark makers, or eyeglass producers.<sup>35</sup> In similar ways, a searchable, online, full-text index is a supplement to the book market (and book culture), not a substitute for it. The difference, of course, is that Google must make copies to accomplish this task, thus confronting the fundamental right to copy that the

---

<sup>31</sup> See Press Release, Ass’n of Am. Publishers, Google Library Project Raises Serious Questions for Publishers and Authors (Aug. 12, 2005), *available at* <http://www.publishers.org/press/releases.cfm?PressReleaseArticleID=274>.

<sup>32</sup> For an argument that Google’s Library Project cannot help but promote book sales, see Posting of Cory Doctorow, Why Publishing Should Send Fruit-Baskets to Google, [http://www.boingboing.net/2006/02/14/why\\_publishing\\_shoul.html](http://www.boingboing.net/2006/02/14/why_publishing_shoul.html) (Feb. 14, 2006, 05:03:24 AM). For questions and doubts about the quality and effectiveness of Google’s book search service in general, see posting of Siva Vaidhyanathan, The Great Unanswered Question: Can Google Do It Right?, <http://www.nyu.edu/classes/siva/archives/002811.html> (Feb. 20, 2006, 07:16 PM).

<sup>33</sup> See Press Release, Ass’n of Am. Publishers, Publishers Sue Google Over Plans To Digitize Books (Oct. 19, 2005), *available at* <http://www.publishers.org/press/releases.cfm?PressReleaseArticleID=292>.

<sup>34</sup> See, e.g., *Lexmark Int’l, Inc. v. Static Control Components, Inc.*, 387 F.3d 522 (6th Cir. 2004) (concluding that section 1201 of Digital Millennium Copyright Act does not apply to potential uses that merely attempt to tether secondary goods like toner cartridges to printers).

<sup>35</sup> Correspondence from Cory Doctorow, blogger, journalist, and author, to Siva Vaidhyanathan, Assoc. Professor of Culture and Cmty., N.Y. Univ. (on file with author).

---

---

Copyright Act grants to holders.<sup>36</sup> So while the publishers' complaints are specious and overreaching, they might just have the law on their side. The conflict over Google's Library Project raises questions that get to the heart of copyright:

- (1) What matters more, the rights of established authors and firms or those firms that push bold and radical plans that free-ride on the work of others?
- (2) What matters more, the fact that Google would make unauthorized copies of entire works for their own cache (an apparent violation of the exclusive rights reserved for copyright holders under the Copyright Act) or the interface that users will experience (a reliance on the fair use provisions of the Copyright Act)?
- (3) Have the processes of digitization and networking altered the assumptions of copyright, demanding a flexible and open vision of copyright?
- (4) Will copyright cease to exist as a copy right and instead morph into a commercial distribution right, as scholars such as Jessica Litman have suggested?<sup>37</sup>

Google's latest venture into the world of print offers us at least three reasons to worry about its bold move to serve as a major — if not the chief — source for book-based information searches: privacy, privatization, and property. The controversies surrounding privacy contribute to the overall public good that Google provides. The issue of privatization speaks to qualitative issues: would Google's service be better for the public than a slower and more piecemeal process of digital indexing governed by libraries and the principles, values, and ethical codes of librarians? Because Google's defense in the property realm depends on convincing courts that its service is in the public interest, both privacy and privatization matter to the property debate, not to mention the overall policy questions concerning Google's emerging role as a dominant force in our information ecosystem.

---

<sup>36</sup> Copyright Act, 17 U.S.C. § 106 (2006).

<sup>37</sup> See Jessica Litman, *Revising Copyright Law for the Information Age*, 75 OR. L. REV. 19, 44 (1996). As Litman wrote, "In addition to separating copyright owners from a useful tool for overreaching, abandoning the reproduction right in favor of a right of commercial exploitation would have the benefit of conforming the law more closely to popular expectations. That would ease enforcement, and make mass education about the benefits of intellectual property law more appealing." *Id.*

*A. Privacy*

Privacy has been a problem for Google, or more precisely, for Google users, for some time. Scores of newspaper and magazine articles have considered the complications of finding one's personal history or long-lost sappy poems accessible via Google.<sup>38</sup> With the launch of Gmail, it became instantly clear that Google's machines were reading our mail for hints about how it might target ads to us.<sup>39</sup> Plainly stated, Google keeps detailed dossiers on the web search habits of its customers. Such information reveals (or falsely indicates) predilections and preferences of millions of people. Such information is valuable for targeting advertising — Google's core revenue. But it is also tempting for state security services to subpoena Google's search records to engage in profiling or surveillance of individuals or groups. Google's policy of retaining such data is deeply troubling. Yet the company gives no indication that it will change its policy, despite loud calls for it to purge its files of identifying information.<sup>40</sup>

Scholars have recognized for some time that in addition to the astounding increase in access to information that the Internet affords, personal information gets swept up in the public stuff. The unregulated flows of, and markets for, personal information have generated much concern. The copyright system has become a surveillance system on the Internet because, in the interest of enforcing copyright, institutions have devised a range of technologies to monitor and govern use of copyrighted material. Apple's ability to monitor the use of iTunes is the best example of this.<sup>41</sup> The repercussions for free, open, playful, creative, and marginal thought are potentially dire.<sup>42</sup>

---

<sup>38</sup> See, e.g., *Google's Long Memory Stirs Privacy Concerns*, CNN.COM, June 3 2005, available at [http://www.cnn.com/aol/story/2005/06/03/cnn\\_internet\\_google.privacy.reut.html](http://www.cnn.com/aol/story/2005/06/03/cnn_internet_google.privacy.reut.html).

<sup>39</sup> See [GoogleWatch.org](http://www.googlewatch.org), Presumably You Have a Gmail Account, and Do Not Object to Google's Policies: But Many of Us Will Not Send Mail to Gmail.com, <http://www.google-watch.org/gmail.html> (last visited Feb. 22, 2007).

<sup>40</sup> See, e.g., Tim Wu, *Keeping Secrets: A Simple Prescription for Keeping Google's Records out of Government Hands*, SLATE, Jan. 23, 2006, <http://www.slate.com/id/2134670/>.

<sup>41</sup> See David Shenk, *A Growing Web of Watchers Builds a Surveillance Society*, N.Y. TIMES, Jan. 25, 2006, at G6.

<sup>42</sup> See generally Oscar H. Gandy, *THE PANOPTIC SORT: A POLITICAL ECONOMY OF PERSONAL INFORMATION* (1993) (describing cultural and political effects of massive collection and collation of personal information by powerful institutions); Julie E. Cohen, *A Right to Read Anonymously: A Closer Look At "Copyright Management" In Cyberspace*, 28 CONN. L. REV. 981 (1996) (describing surveillance powers granted to

With Google Library, we have a whole new set of privacy concerns. How will we be able to trust Google not to turn over individual patron reading records to the FBI or local law enforcement officials on fishing expeditions? There is nothing in Google's privacy policy that promises it will resist such abusive practices.<sup>43</sup> Besides, plenty of other firms, including airlines, have violated their own privacy policies to do just that. More importantly, nothing in the contract that outlines the terms for the Google Library Project between the University of Michigan and Google binds Google to respect patron confidentiality in any way.<sup>44</sup>

### B. Privatization

Privatization of library functions not necessarily a bad thing. We should not pretend that libraries operate outside market forces or do not depend on outsourcing many of their functions to private firms. But we must recognize that many of the thorniest problems facing libraries today are a direct result of rapid privatization and onerous contract terms. Onerous contract terms with vendors of electronic material severely limit both the abilities librarians have to serve their patrons and the fair use rights of both librarians and patrons.<sup>45</sup> There are too many devils in too many details.

It is important to remember that Google serves its own masters: its stockholders and its partners. It does not serve the people of the State of Michigan or the students and faculty of Harvard University. The real risk of privatization is simple: companies fail. Libraries and universities last. Companies wither and crash. Should we entrust our heritage and collective knowledge to a company that has been around for less than a decade? What if stockholders decide that Google Library is a money loser or too much of a copyright liability? What if

---

copyright holders by Digital Millennium Copyright Act and digital rights management tools); Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373 (2000) (arguing that privacy debates should be grounded in considerations of "the conditions necessary for individuals to develop and exercise autonomy in fact, and that meaningful autonomy requires a degree of freedom from monitoring, scrutiny, and categorization by others").

<sup>43</sup> See Google, Google Privacy Center, <http://books.google.com/privacy.html> (last visited Feb. 22, 2007).

<sup>44</sup> See Cooperative Agreement, *supra* note 29.

<sup>45</sup> Communication from James Neal, Director of Libraries, Columbia Univ., to Siva Vaidhyanathan, Assoc. Professor of Culture and Cmty., N.Y. Univ. (on file with author).

they decide that the infrastructure costs of keeping all those files on all those servers do not justify the expense? What then?

The celebration of Google's Library Project reveals a dangerous assumption: that the role of the librarian in the global digital information ecosystem is superfluous. It also ignores serious qualitative issues. Google has neither revealed nor discussed the principles upon which the search engine will operate. In contrast, librarians and libraries operate with open and public standards for metadata and organization. Metadata is particularly important in this question.<sup>46</sup> Full-text searching is an insufficient and often absurd method of seeking information. Without metadata — data about data — embedded in the files to guide search engines via subject headings, keywords, and quality indicators, a search for books about the Holocaust is just as likely to reveal books denying the event as examining it. Metadata can guide a computer-based search to favor certain sources over others. Good metadata standards generate better searches. Poor metadata standards might yield dangerous or ridiculous results.<sup>47</sup> So far, we have no reason to believe that the privatization of this indexing function from a public university library to a private entity will involve good or open metadata standards. The contract between Michigan and Google demands no such quality control.<sup>48</sup>

### C. *Property*

Let me be blunt: Google Library invites copyright meltdown. If all goes as Google plans, we might not have a copyright system we recognize in ten years. And that is not necessarily a good thing.

Paul Ganley, a London-based solicitor, has written an analysis of the Google library case under both U.S. and U.K. law. He concludes that while Google has a slight chance of prevailing under the flexible fair use provisions of U.S. law, it has absolutely no chance of surviving a challenge in U.K. courts. Ganley uses the case as a “teaching moment” because to him it generates two wonderful potential exam

---

<sup>46</sup> See generally Ganesan Shankaranarayanan & Adir Evan, *The Metadata Enigma*, 49 *COMMS. ACM* 88 (2006) (describing methods and utility of metadata assignment and analysis).

<sup>47</sup> See generally Lucas D. Intronza & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matter*, 16 *INFO. SOC'Y* 1 (2000).

<sup>48</sup> See Cooperative Agreement, *supra* note 29.

questions: can Google do this under existing copyright law? Should Google be able to do this under copyright law?<sup>49</sup>

I would add a third question that speaks directly to the first two: is Google the right agent to do this? If Google is the right agent, then copyright law certainly should allow ambitious and potentially beneficial uses of copyrighted material that on balance do not threaten existing markets for works. However, if it is not the best institution, if some other institution should be taking the risk and reaping the rewards for such a project, perhaps copyright law already allows other institutions better suited for these efforts to undertake them.

I conclude that legally, politically, and practically, Google is not the right agent for the job. Instead, libraries should be pooling their efforts and resources to accomplish such massive digitization and access projects. And because Google is so inappropriate, its legal argument is inherently weakened, thus answering Ganley's first question negatively.

## II. THE ARGUMENTS FOR GOOGLE

Distinguished scholars and litigators such as Jonathan Band,<sup>50</sup> William Patry,<sup>51</sup> Fred von Lohmann,<sup>52</sup> Cory Doctorow,<sup>53</sup> and Lawrence Lessig<sup>54</sup> have all voiced enthusiasm for the Google Library Project and launched defenses of the firm's copyright strategy. Each of these writers relies on the traditional (and statutory) "four factor" fair use

---

<sup>49</sup> See Paul Ganley, *Google Book Search: Fair Use, Fair Dealing, and the Case for Intermediary Copying* 1, 1 (Jan. 13, 2006) (unpublished manuscript, on file with the Social Science Research Network), available at [http://papers.ssrn.com/so13/papers.cfm?abstract\\_id=875384](http://papers.ssrn.com/so13/papers.cfm?abstract_id=875384).

<sup>50</sup> See generally Jonathan Band, *The Google Library Project: Both Sides of the Story*, 2 *PLAGIARY* 1 (2006), available at <http://www.plagiary.org/Google-Library-Project.pdf> (laying out fair-use analysis of Google service).

<sup>51</sup> See generally Posting of William Patry, *Google Revisited*, <http://williampatry.blogspot.com/2005/09/google-revisited.html> (Sept. 23, 2005) (arguing that "snippet" was sufficiently transformative to qualify as fair use of longer work). Patry has since removed this blog posting. An excerpt of the original posting is available at <http://digilib.bu.edu/blogs/theolib/?p=37> (last visited Feb. 14, 2007).

<sup>52</sup> See generally Fred von Lohmann, *Electronic Frontier Found.: Deep Links: Authors Guild Sues Google* (Sept. 20, 2005), <http://www.eff.org/deeplinks/archives/003992.php> (offering "four-factors" analysis of Google project).

<sup>53</sup> See generally Posting of Cory Doctorow, *supra* note 32 (positing that it would be to publishers' and authors' advantage to let Google proceed with library scanning project).

<sup>54</sup> See generally Lawrence Lessig, *digital video: Is Google Book Search Fair Use?* (2006), <http://www.youtube.com/watch?v=512nrBmBQXg> (presenting "market failure" case for Google's fair use claim).

analysis of Google's use of the works.<sup>55</sup> In each case, they minimize the fourth factor, that of "harm to the potential markets" of the original work by declaring that the Google project would not harm the sale of books and might enhance it. In addition, they concur that several important cases in recent years have shown that commercially viable uses are not beyond the scope of fair use.<sup>56</sup> In each argument, they treat the "snippet" of text that Google users would encounter when clicking on a link to a copyrighted work as the operative use of the work. They minimize the importance of the original scanning of the book, the very copying that the publishers want the court to consider as operational and significant. They argue that the snippet-based interface is "transformative," thus invoking the magic word that Justice Souter employed in his ruling for the hip-hop group 2 Live Crew in the case of *Campbell v. Acuff Rose*.<sup>57</sup> By "transforming" the original song, Roy Orbison's classic hit "Oh, Pretty Woman," the defendant, Luther Campbell, created something entirely new — in this case, a parody of the original song. "Transformativeness" stands now as a process distinct from "derivativeness." If a work is derivative of a copyrighted work, it is under the control of the copyright holder. If the work is considered "transformative," it is considered fair use.<sup>58</sup> So there is much at stake in the distinction. When considering the composition of distinct creative works, it serves well to have a broad and strong sense of transformation. But, as Michael Madison points out, there is nothing close to consistency in the ways courts establish the transformativeness of a use.<sup>59</sup>

In addition, the defenders of Google's copyright strategy all rely on the claim that a snippet would obviously be but a small portion of a book and, thus, of a work. Thus, they aid Google in the consideration of the third factor: the amount and substantiality of the taking. The

---

<sup>55</sup> See Copyright Act, 17 U.S.C. § 107 (2006). This section instructs courts to consider four factors when judging whether an unauthorized use of copyrighted material is "fair" and, thus, non-infringing. 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." *Id.*

<sup>56</sup> See, e.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 585 (1994); *Ty, Inc. v. Publ'ns Int'l, Ltd.*, 292 F.3d 512, 515, 523 (7th Cir. 2002).

<sup>57</sup> See *Campbell*, 510 U.S. at 578-79.

<sup>58</sup> *Id.* at 579.

<sup>59</sup> Michael J. Madison, *A Pattern-Oriented Approach to Fair Use*, 45 WM. & MARY L. REV. 1525, 1670-71 (2004).

problem with this assertion, of course, is that often books are composed of small, distinct works of authorship — an anthology of poetry, for instance. A standard four-factor fair use analysis of the “snippetization” of Leo Tolstoy’s *War and Peace* would look very different from one concerning a collection of haiku or limericks. For these and other reasons, the pedestrian exercise and almost arbitrary nature of the four-factor test has driven many scholars and judges to question its utility.<sup>60</sup>

Lessig bolsters his argument by invoking an economic theory of fair use: that fair use is justified by market failures.<sup>61</sup> If a market cannot supply an important public good or service, it has failed. Market failures can be invoked to justify state intervention in a process to ensure the production or distribution of the good or service. Or, as in this case, state inaction (or the negation of a state-granted right) could be justified on the basis of market failure. But the important thing about market failures is that the good or service in question must be beneficial and significant. If a service’s effects are trivial or deleterious, it cannot qualify as a market failure. Market failures within copyright are quite common. For years, copyright maximalists, such as Bruce Lehman, have been arguing that fair use can be displaced in the digital environment because users and copyright holders can contract to use small portions of work with the smallest of transaction costs.<sup>62</sup> Monitoring of digital uses facilitates micro-payments for many of the uses that used to be considered “fair.” If a teacher wants to hand out forty copies of a story from that day’s newspaper to her class, she has the ability to clear the distribution and pay a very small negotiated rate to the newspaper using digital tollbooths and other similar payment systems. Clearly, those of us who believe that fair use rests on republican principles, that fair use generates a penumbra of users’ rights and a level of comfort and confidence, should recoil at such economically based theories of fair

---

<sup>60</sup> See, e.g., *Ty*, 292 F.3d at 521; *Madison*, *supra* note 59, at 1530; Georgia Harper, Univ. of Tex. Sys., Office of the Gen. Counsel, *Google This* (Oct. 19, 2005), <http://www.utsystem.edu/ogc/INTELLECTUALPROPERTY/googlethis.htm>.

<sup>61</sup> See Lessig, *supra* note 54. For an example of an argument asserting that fair use is justified by market failures, see generally Wendy Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600 (1982).

<sup>62</sup> See BRUCE A. LEHMAN, U.S. EMBASSY IN THAILAND, PROTECTING INTELLECTUAL PROPERTY IN A GLOBAL ECONOMY, <http://bangkok.usembassy.gov/services/docs/reports/ipr1.htm> (last visited Feb. 14, 2007); see also Pamela Samuelson, *The Copyright Grab*, WIRED, Jan. 1996, available at [http://www.wired.com/wired/archive/4.01/white.paper\\_pr.html](http://www.wired.com/wired/archive/4.01/white.paper_pr.html).

use. Yet Lessig invokes just such a theory. He argues that Google cannot clear the rights to the great majority of books published since 1923 because there is no way to determine who controls the rights. This is a classic market failure: to facilitate this essential public good — widespread dissemination of book learning — courts should grant Google broad fair use rights to scan in all of the millions of books without permission.<sup>63</sup> It is a compelling argument. If we want such a service and we deem it valuable to the general good, clearly the institution that pursues it should not have to seek permission to engage in the basic scanning of the works.

### III. PRAGMATISM OR FUNDAMENTALISM?

All of these defenders of Google rest their argument on the important precedent established in *Kelly v. Arriba Soft Corp.* In *Kelly*, the California-based Ninth Circuit Court of Appeals made a pragmatic ruling. The court declared the thumbnail image of a cached image sufficiently transformative to qualify as non-infringing under *Campbell*.<sup>64</sup> Mostly, the *Kelly* ruling clearly demonstrated that effective search engines are essential to the culture and economy of the Internet and the United States. The court injected flexibility and realism into the copyright system, making it fit the realities of both creative and economic practices.

But over on the East Coast, the Second Circuit Court of Appeals has been much less pragmatic. Faced with a similar situation, the Second Circuit ruled against a company called MP3.com that wanted to rip music from lawfully purchased compact discs and place the files in a protected “digital locker” so that subscribers could access their own music (they had to prove they had purchased the discs themselves). A subscriber would log on to the MyMP3.com site and gain access to MP3s of the exact catalog that she had purchased in CD form. To facilitate this, of course, MP3.com had to create a digital archive of thousand of discs. This service did not harm the market for compact discs in any way. In fact, music companies had to sell at least two discs for this transaction to occur. Nonetheless, the District Court for the Southern District of New York ruled fundamentally: MP3.com did not have permission to copy and stream in their entirety the compact discs it had purchased.<sup>65</sup>

---

<sup>63</sup> See Lessig, *supra* note 54.

<sup>64</sup> *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818-22 (9th Cir. 2003).

<sup>65</sup> *UMG Recordings, Inc. v. MP3.com, Inc.*, 92 F. Supp. 2d 349, 352-53 (S.D.N.Y. 2000).

The Second Circuit and the Southern District of New York have issued other fundamentalist rulings in matters of copyright.<sup>66</sup> Although it is not a relevant legal precedent to the Google case, the most revealing of these is *The New York Times Co. v. Tasini*. The rhetoric of this case tracks the Google library debate quite presciently. The *New York Times* and other publications neglected to secure digital rights to thousands of articles submitted by freelance writers. The Writers' Union sued because the *Times* and others were selling electronic copies of their works to research databases such as Lexis-Nexis and ProQuest. The Writers' Union argued straightforwardly that without expressed permission, the periodicals were violating the freelancers' rights to control distribution and derivative works. The periodicals countered that they were acting on an implied license and that the inclusion of these works in these popular and valuable databases served the public interest. Both the Second Circuit and the Supreme Court sided with the writers. They dismissed all "market failure" and "public good" arguments. They instead opted for a fundamental reading of copyright.<sup>67</sup>

#### IV. A LAPSE INTO LEGAL REALISM

As Lawrence Lessig has said and written on many occasions, "fair use is the right to hire a lawyer."<sup>68</sup> This statement is a recognition of the central problem of fair use, its central paradox: while fair use might seem to be growing stronger on paper (and in court),<sup>69</sup> it is

---

<sup>66</sup> See, e.g., *Am. Geophysical Union v. Texaco, Inc.*, 37 F.3d 881 (2d Cir. 1994) (ruling that research compiled in private corporate library does not enjoy broad fair use protection); *Universal City Studios, Inc. v. Reimerdes*, 111 F. Supp. 2d 346 (S.D.N.Y. 2000) (ruling that journalists are not exempt from prohibitions against distribution of circumvention technologies banned by section 1201 of Digital Millennium Copyright Act).

<sup>67</sup> See *Tasini v. N.Y. Times, Co.*, 206 F.3d 161 (2d Cir. 2000), *aff'd*, 533 U.S. 483, 502-05 (2001) (ruling that without explicit contract allowing *The New York Times* to repackage and sell freelance articles acquired under contracts signed before advent of digital library services such as Lexis-Nexis, *The New York Times* had infringed on rights of freelance authors).

<sup>68</sup> Posting of Lawrence Lessig to Lessig Blog, <http://www.lessig.org/blog/archives/001794.shtml> (Mar. 20, 2004, 18:26 EST).

<sup>69</sup> See, e.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994) (expanding fair use to cover parodies); *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984) (including private noncommercial uses such as "time-shifting" as fair uses); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003) (ruling that thumbnail images are fair use abstractions of larger copyrights images); *Ty, Inc. v. Publ'ns Int'l Ltd.*, 292 F.3d 512 (7th Cir. 2002) (including some catalogs as protected by fair use despite their clear commercial nature); *Suntrust Bank v. Houghton Mifflin Co.*, 136 F. Supp.

increasingly less fair and less useful in real life. Not every publicly beneficial use is a parody or a thumbnail. The confidence that fair use affords creators correlates strongly with one's position in the socioeconomic scale and one's expertise in matters of copyright. Google can hire a whole lot of lawyers. Google may suffer from a surplus of confidence. To get a sense of the magnitude of this confidence, compare Google's optimistic claim that it can win a court battle for its Library Project, with the misfortunes of a smaller, poorer company called Video Pipeline. A New Jersey-based company, Video Pipeline had for years supplied taped video "clip" catalogues of Hollywood films to movie rental stores and vendors. Such a service allowed vendors to preview and choose the works they would select for their stores. In the late 1990s, just as Video Pipeline was developing an online version of its catalogue service, the Disney Corporation decided it no longer wanted Video Pipeline to perform such a service. Disney could do its own video catalogues and marketing. Video Pipeline continued to include Disney films among the snippets it offered in its market. Disney sued, pointing out that Video Pipeline's practices violated several of its rights under copyright, including distribution, public performance, and the production of derivative works. Video Pipeline countered that its use of small portions of these films did not retard the market for the films and videos. If anything, it enhanced it. Furthermore, the snippets were so small that they must qualify as fair use. The Third Circuit Court of Appeals disagreed with Video Pipeline, sided with Disney, and effectively shut down the small company.<sup>70</sup>

#### V. WHY QUALITATIVE FACTORS MATTER

Lessig's defense of Google depends on courts agreeing with his assertion that the Google Library Project "could be the most important contribution to the spread of knowledge since Jefferson dreamed of national libraries. It is an astonishing opportunity to revive our cultural past, and make it accessible."<sup>71</sup> Such hyperbole is essential to Google's case. If the Google Library Project does not promise to deliver a substantially valuable research tool then the market failure is

---

2d 1357 (N.D. Ga. 2001) (expanding the parody exception further).

<sup>70</sup> For information on the circumstances leading up to this lawsuit, as well as the court's ruling, see generally *Video Pipeline, Inc. v. Buena Vista Home Entm't*, 342 F.3d 191 (3d Cir. 2003).

<sup>71</sup> Posting of Lawrence Lessig to Lessig Blog, *Google Sued*, <http://lessig.org/blog/archives/2005-09.shtml> (Sept. 22, 2005 5:37 A.M.).

irrelevant. A useless research tool would not deserve a fair use exemption. The problem with Lessig's argument is that Google's search algorithms, while effective (yet imperfect) for something as dynamic and ephemeral as the Web, are wholly inappropriate for stable texts like books. Any simple search of terms like "It was the best of times" or "copyright" will yield very bad results. Google will have a hard time convincing a court that we are actually better off with this service than without it. The fundamental error that Google and many of its supporters make is assuming that Google's algorithms and selections are somehow neutral, that they do not betray certain biases in them. The prevalence of computer manuals in Google Book Search searches betrays just one of the biases built into the software. There are many more yet to be discovered. We must start our consideration of Google Book Search from the principle that no technology is neutral.<sup>72</sup> But beyond neutrality, the fact is that Google's PageRank algorithm is just good enough for the web. It is hardly an effective or comprehensive research tool. It generates too many ridiculous results for simple searches. It cannot screen out bad results very well. And Google offers no simple information-seeking training to its customers. Searching the text of books is rarely a better way to search than searching among books. Books are discreet documents that operate with internal cohesion more than external linkages. They are not "small pieces loosely joined,"<sup>73</sup> nor should they be. Their value is in their comprehensiveness. Printed and bound books are examples of a portable, reliable technology that has worked extremely well for more

---

<sup>72</sup> See, e.g., Michael T. Zimmer, *Media Ecology and Value Sensitive Design: A Combined Approach to Understanding the Biases of Media Technology* 1, 3 (2005) (unpublished Ph.D. dissertation, New York University) (on file with the author) (arguing that presumption of technological neutrality blinds us to biases designed into those technologies themselves).

<sup>73</sup> See generally DAVID WEINBERGER, *SMALL PIECES, LOOSELY JOINED: A UNIFIED THEORY OF THE WEB* (2002) (describing how Internet has created "a loose federation of documents — many small pieces loosely joined"). Weinberger notes that the Internet has changed not just the way documents are connected to one another, but has also changed the interior structure of documents:

The Web has blown documents apart. It treats tightly bound volumes like a collection of ideas — none longer than can fit on a single screen — that the reader can consult in the order she or he wants, regardless of the author's intentions. It makes links beyond the document's covers an integral part of every document. What once was literally a tightly-bound entity has been ripped into pieces and thrown into the air.

*Id.*

than five hundred years. No one has yet shown that searches results of “key words in context” have much value to readers, researches, or writers. Privileging textual searching over more established forms of book indexing is a mistake. Relying on Google’s engineers to do the work that librarians do is a bigger mistake.

#### VI. THE TRUMP CARD: THE LIBRARY COPY

Although they have yet to file suit over the scanning of their books, representatives of American university presses have a particular complaint against Google and the University of Michigan library. They say that there was already an emerging and mutually beneficial market for electronic access to, and indexes for, books that academic presses would provide to libraries. For more than a decade, many university presses have been formatting and storing electronic files of back-catalogue and out-of-print work. Some of these were intended to generate a “print-on-demand” market for books for which demand was too limited to justify print runs of several hundred copies per year. But with the support of foundations, university presses were engaged in forming a consortium that would standardize the format and index and offer electronic books to libraries for a subscription fee. Since the announcement of the Google Library Project, the many academic publishers and libraries have suspended such projects.<sup>74</sup> The Google project has the potential to crowd out many experiments and initiatives. The university press directors are not particularly troubled by the open Web search capabilities that Google offers. Instead they wonder about the propriety and legality of the transfer of the electronic copy from Google back to the University of Michigan library.<sup>75</sup>

We should be troubled as well. I have asked many scholars and activists who support Google’s position on this project: what possible justification under fair use or any other provision or exemption under copyright is there for the creation and distribution of an entire

---

<sup>74</sup> For the sake of honesty and transparency, I must disclose that I served as a paid consultant for such a consortium organized by Oxford University Press in 2004. The project ended abruptly when its leaders learned of Google’s plans to undermine its potential market. Oxford University Press paid me a one-time fee of \$1,000 before the project folded. I did not expect any subsequent compensation regardless of the prospects of the project. I did, however, support the aims of the project before I signed on as a consultant.

<sup>75</sup> Correspondence from Steven Maikowski, Director, N.Y. Univ. Press, and Niko Pfund, Academic Publ’g Director, Oxford Univ. Press — U.S., to Siva Vaidhyanathan, Assoc. Professor of Culture and Cmty., N.Y. Univ. (on file with the author).

copyrighted work as payment for a commercial transaction? I know of no such principle, no such precedent, and no such exemption. I have yet to receive an answer.

I submit that even if — in the unlikely event that a court in the Southern District of New York and the Second Circuit Court of Appeals, and the Supreme Court were all to ignore their own guiding precedents and instead expand the scope of a fragile Ninth Circuit precedent — Google wins its argument that the user interface matters more than the trivial archiving of the entire work, Google cannot win the case at large. The “payback copy” that would go to the University of Michigan library is Google’s fatal weakness.

#### CONCLUSION

I am not a litigator. I am not an advocate. I am not even a lawyer. But I am a researcher, writer, author, and Web enthusiast. I have been calling for a more open and flexible copyright system for more than a decade. I celebrate the importance of such cases as *Kelly* and *Sony Corp. v. Universal City Studios, Inc.*, both of which cleared space for corporate exploitation of others’ copyrighted works for the greater good. But I fear the Google Library Project is the wrong cause. It is destined to backfire on the emerging “Free Culture” movement. It directly jeopardizes the decision in *Kelly*. If a court issues an indelicate, clumsy, or violently sweeping ruling in the pending Google cases concerning publishers and authors, we could lose many important public rights under copyright, including special library privileges. Even if Google prevails, we are risking the stability of the library system, because Congress, always ready to do the bidding of the copyright industries, certainly would exact revenge on Google and its academic partners for this hubris by revising the Copyright Act to redress what the publishers would see as a significant and unwarranted limitation on the exclusive rights of copyright holders. If any court either overturns or significantly narrows the scope of *Kelly*, the entire Internet (and everyone who uses it) will pay the price. Some users would express frustration that we have a copyright system that inhibits basic research and communication and that runs so contrary to common sense. Others would make do by eschewing the Web’s radically democratic nature and immersing themselves in a more corrupt, ordered, and top-down information world in which contracts, payments, and permissions guide searches and rankings. The copyright system would have failed us again. The only hope to revive an open and democratic Web would include a movement to adopt formalities in copyright once again, so that search engines

would not have to presume they had a license or negotiate for one for every Web page in the universe. The prospects for that are slimmer than that of the Second Circuit overturning itself in *MP3.com* and *Tasini*. Not so long ago we thought the copyright meltdown would come from music. Who knew it would come from books?