

## **Google Book Search – Fair Use or Infringement?**

As we all know, Google has started to take over. With everything from email and search services to a Web-based word processor, the world is becoming “Googlized.” Therefore, it is not surprising that, at times, there will be both critics and proponents debating the ethical and legal standards surrounding some of Google’s products. The Google Book Search project is an example that raises both moral and lawful arguments, while testing the mature concept of fair use in the current digital era.

### **Background**

In 2004, Google announced a new program, Google Print (which was later renamed Google Book Search), in which the company scans books, converts them to text using optical character recognition, and stores the books in a digital database making the full text of these books searchable.

When relevant to a user's keyword search, up to three results from the Google Book Search index are displayed above search results in the Google Web Search service (google.com). A user may also search just for books at the dedicated Google Book Search service Web page. Clicking a result from Google Book Search opens an interface in which the user may view pages from the book as well as content-related advertisements and links to the publisher's Web site and booksellers. Through a variety of access limitations and security measures, some based on user-tracking,

Google limits the number of viewable pages to 20 percent of the text and attempts to prevent page printing and text copying of material under copyright.

Additionally, Google allows rights holders to proactively “opt out” of the scanning (a point of debate for critics and proponents), rather than only scanning copyrighted materials for which the rights holders have given their explicit permission to be included.

The Google Book Search function has continued to expand since its inception in 2004. Originally, Harvard, Stanford, Oxford, The New York Public Library and the University of Michigan were the only libraries that agreed to the scanning of their book collections. Since, this list has expanded to various other university and public libraries, as well as a number of magazines, such as New York Magazine and Popular Mechanics. Additionally, Google Book Search now includes books in English, French, Italian, German, Spanish, Latin and Dutch. And, the service currently allows public-domain works and other out-of-copyright material to be downloaded in PDF format.

In October 2009, Google announced that the number of scanned books had reached to over 10 million. And, as of right now, the texts in Google Books come from two sources:

- The Library Project – As mentioned above, Google has partnered with renowned libraries around the world to include their collections in Book Search. For the Library Project books that are still in copyright, the Google Books Search results are like a card catalog; showing the consumer information about the book and, generally, a few snippets of text showing the

search term in context. For Library Project books that are out of copyright, however, the consumer can read and download the entire book.

- The Partner Program – Google has also partnered with over 20,000 publishers and authors to make their books discoverable on Google. Consumers can flip through a few preview pages of the books, just like browsing them at a bookstore or library. Readers also see links to libraries and bookstores where one can borrow or buy the book.

### **Dueling Viewpoints**

In terms of the legal and ethical issues surrounding the Google Book Search project, there are two clear opposing viewpoints, both of which have strong cases. Overall, the initiative has been hailed for its potential to offer unprecedented access to what may become the largest online corpus of human knowledge and promoting the democratization of information, but it has also been criticized for potential copyright violations.

Those in support of the Google Book Search project believe it makes finding and previewing rare and out-of-print books easier – this includes foreign language books that might not typically be readily available. Proponents feel strongly that it is beneficial to provide U.S. users with a way to view entire texts of scanned books at public or university libraries, and at schools that have requested an “institutional subscription,” because it delivers wider access to books for everyone, especially students. Proponents also believe that if Google Book Search were able to expand its offerings, more book text would become searchable over the Internet. And,

considering authors would get compensated for the usage, they would gain an additional method of selling their work – authors and publishers would receive payments when users read their books online. Out-of-print books still under copyright would fall under the same royalty system as in-print books, giving authors a previously nonexistent market for making money off their no-longer-published titles.

In contrast, critics of the service feel that Google could gain an unfair ability to profit from books and believe that, as the sole Internet book search service of its sort, Google could theoretically raise prices without any competition to keep it under control. Additionally, challengers believe that by having the expanded searchable book content within its own servers, Google could have an unfair edge over its search engine competitors and could also have a negative effect on other book sellers and distributors, creating such a large system that they will be unable to compete. Authors, as well as publishers, feel that the service (and the pending settlement) would give Google too broad a license over their work. Even with the revenue-sharing system in place, many authors feel Google would be stealing some of their rights and question the effectiveness of an "opt-out" option.

One other area of concern is a type of book known as an "orphan work" – a book for which the rights-holders are not known, or the author cannot be located. Critics say Google's Book Search deal would give Google an exclusive license and full control over these titles, allowing it to distribute and price them at its own discretion.

## **The Battle**

With these concerns in hand, three years ago the Authors Guild, the Association of American Publishers (AAP) and a handful of authors and publishers filed a class action lawsuit against Google Books, challenging Google's digitization and dissemination of in-copyright books without explicit permission from rightsholders.

The authors and publishers felt Google infringed the rights of copyright holders when it scanned entire books and stored the digitized versions in its massive database.

Google argued that creating an easy-to-use index of books was considered fair use under copyright law.

### **Understanding Fair Use**

Fair use is a doctrine in United States copyright law that allows limited use of copyrighted material without requiring permission from the rightsholders.

The origins of fair use date back to the legal concept of "test copyright," which was first ratified by the Kingdom of Great Britain's Statute of Anne of 1709. Within this newly formulated statutory right, room was not made for the authorized reproduction of copyrighted content; therefore, the courts created a doctrine of "fair abridgement" in the *Gyles v. Wilcox* case (1869), which later eventually evolved into the modern concept of "fair use" that recognized the utility of such actions.

In the case, Lawrence sued Dana et al. for infringing on Lawrence's copyrights.

Lawrence had edited and annotated two volumes of *Wheaton's Elements of International Law*, memoirs of the deceased Mr. Wheaton, upon the request of the

author's wife, Mrs. Wheaton. At the time of the lawsuit, Mrs. Wheaton had died and a different publisher had published the memoirs with no accounting to Lawrence. The court noted that Lawrence's annotations and editing "involved great research and labor" and ultimately found the use of Lawrence's materials not to be fair. While Dana et al. argued that they had merely abridged, or fairly used, Lawrence's materials, the court said the use was far more than a fair use; it was a reprint.

The fair use doctrine only existed in the U.S. as common law until it was incorporated into the Copyright Act of 1976, Section 107.

Section 107 contains a list of the various purposes for which the reproduction of a particular work may be considered fair, such as criticism, comment, news reporting, teaching, scholarship, and research. Section 107 also sets out four factors to be considered in determining whether or not a particular use is fair:

1. The purpose and character of the use, including whether such use is of 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
4. The effect of the use upon the potential market for, or value of, the copyrighted work

The four factors of analysis for fair use set forth derive from the 1841 case of *Folsom v. Marsh*. In *Folsom*, the defendant wrote a biography of George Washington but

used 353 pages of the plaintiff's earlier published and copyrighted multivolume work to do so. Although the defendant's use amounted to less than six-percent of the plaintiff's total work, the court held for the plaintiff, finding the defendant had copied the most important material (i.e., substantiality) in the plaintiff's earlier volumes.

One thing to consider is that the distinction between fair use and infringement may be unclear and not easily defined. There is no specific number of words, lines or notes that may safely be taken without permission.

“The practical effect of this law and the court decisions following it is that it is usually possible to quote from a copyrighted work in order to criticize or comment upon it, teach students about it, and possibly for other uses. Certain well-established uses cause few problems. A teacher who prints a few copies of a poem to illustrate a technique will have no problem on all four of the above factors (except possibly on amount and substantiality), but some cases are not so clear. All the factors are considered and balanced in each case: a book reviewer who quotes a paragraph as an example of the author's style will probably fall under fair use even though he may sell his review commercially. But a non-profit educational website that reproduces whole articles from technical magazines will probably be found to infringe if the publisher can demonstrate that the website affects the market for the magazine, even though the website itself is non-commercial” (“Fair use”).

Additionally, this distinction between fair use and infringement is even harder to understand and control in a digital world. As noted by Lawrence Lessig in *Free*

*Culture*, “For while it might be obvious in the world before the Internet, copies were the obvious trigger for copyright law, upon reflection, it should be obvious that in the world with the Internet, copies should not be the trigger for copyright. More precisely, they should not *always* be the trigger for copyright law” (140). With the Internet, every use of a copyrighted work produces a copy, which allows for previously unregulated uses to be regulated. “A thin protection grounded in fair use makes sense when the vast majority of uses are unregulated. But, when everything becomes presumptively regulated, then the protections of fair use are not enough” (Lessig, *Free Culture* 145).

It is important to note, however, that the uncertainty around the fair use doctrine serves an important purpose in light of fast-changing technologies and globalization. “It leaves space for change; it offers a method whereby the law, so tied to the alphabetic text, can keep up with digital technologies. The downside of uncertainty is that judges struggle with case-by-case, fact-specific analyses, sometimes applying the doctrine inconsistently” and continually creating confusion, as in the Google Book Search situation (Rife).

### **The Settlement**

Following a three-year battle over Google's right to display copyrighted books on its Web site, the Book Search settlement was announced in October 2009. Google agreed to pay \$125 million, in addition to other terms, to ensure authors and publishers could register to receive payments anytime their books were viewed within the service.

In exchange, the Authors Guild and APA dropped their lawsuits, allowing Google to move forward with the project.

Additionally, the settlement clarified that Google planned to work closely with authors, publishers and copyright holders to bring even more of the world's books online.

On November 19, 2009 the court granted preliminary approval for the amended settlement in which the following changes are expected:

- The agreement will allow Google and its publishing industry partners to greatly expand the number of books that readers can find, preview and buy through Google. Out-of-print books will be available for preview, reading and purchase in the U.S.
- The settlement agreement will create new options for reading entire books. Consumers will be able to purchase full online access to millions of books, meaning they can read an entire book from any Internet-connected computer, simply by logging in to their Book Search account. And, the book will remain on the consumer's electronic bookshelf, so he/she can come back and access it whenever desired in the future. Additionally, Google will offer libraries, universities and other organizations the ability to purchase institutional subscriptions, which will give users access to the complete text of millions of titles while compensating authors and publishers for the service. Students and researchers will also have access to an electronic library that combines the collections from many of the top universities across

the country. Additionally, public and university libraries in the U.S. will be able to offer terminals where readers can access the full text of millions of out-of-print books for free. And, Google will continue to point users to nearby libraries and bookstores where the readers' desired books are available.

- Google also agreed to establish a Book Rights Registry to help publishers and authors control how their copyrighted works are accessed. In essence, the Registry will help locate rightsholders and ensure that they receive the money their works earn under the settlement agreement.
- The agreement creates new opportunities for libraries and universities to offer their patrons and students access to millions of books beyond their own collections. In addition to the institutional subscriptions and the free public access terminals, the agreement also creates opportunities for researchers to study the millions of volumes in the Book Search index. Academics will be able to apply through an institution to run computational queries through the index without actually reading individual books.

### **Should Google Book Search be considered fair use?**

Even with the approval of the settlement agreement, the question still remains unanswered: Should Google Book Search be considered fair use?

By settling the lawsuit with book authors and publishers Google avoided the fight for establishing a positive legal precedent for copyright fair use on the Internet. Obviously, both proponents and critics of the service still have their own opinions and copyright experts, such as Lessig, believe Google stood a fair chance of winning in court. As stated by Fred von Lohmann, "It is easy to see how [Google Books] can stimulate demand for books that otherwise would lay undiscovered in library stacks. On the other hand, it is hard to imagine how it could hurt the market for the books—getting a couple sentences surrounding a search term is unlikely to serve as a replacement for the book" ("Google Books"). Consequently, while Google's settlement decision did not test the boundaries of fair use and, more specifically, define the scope of fair use, as it was expected to, the settlement offered an example of how digital companies and copyright holders can compromise in light of the reality that digital content can be so easily and broadly distributed online. "Unlike the music and newspaper industries, the book industry is wise to strike a deal that allows it to benefit from digital distribution without compromising its right. Meanwhile, Google, while not admitting wrongdoing, implicitly acknowledges that publishers and authors have a right to this material. It's a compromise on both sides" (Albanesius).

Therefore, although Google was never taken to court to fully test the current limits of the fair use doctrine, the Google Book Search project created a digital-age test for copyright laws that long preceded the Internet, bringing the idea of fair use to the forefront in a digital society and proving that the Book Search project is a positive step in the right direction regarding increasing our ability to obtain information. As

stated by Lawrence Lessig, "[Google Books] 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" ("Google Books").

## References

- Albanese, Andrew. "Harvard Slams Google Settlement; Others React with Caution." *Library Journal*. 30 Oct. 2008. Web. 1 Oct. 2009.
- Albanesius, Chloe. "DOJ Urges Court to Reject Google Books Settlement." *PC Magazine*. 19 Sept. 2009. Web. 1 Oct. 2009.
- Bray, Hiawatha. "Joyce, Dickens, Google -- classics are there to download." *Boston Globe*. 31 Aug. 2006. Web. 1 Oct. 2009.
- Cornell University Law School. "107. Limitations on exclusive rights: Fair use." *Cornell University Law School*. Web. 25 Nov. 2009.
- "Fair Use." *Wikipedia Online*. Web. 1 Oct. 2009.
- Gonsalves, Antone. "Major Book Publishers Sue Google." *Information Week*. 19 Oct. 2005. Web. 1 Oct. 2009.
- "Google Book Search." *Wikipedia Online*. Web. 1 Oct. 2009.
- Google Books. "Google Books Settlement Agreement." *Google*. Web. 23 Nov. 2009.  
<http://books.google.com/googlebooks/agreement/>
- Jesdanun, Anick. "Google book project: Digital-age test of copyright law." *USA Today*. 18 Sept. 2005. Web. 1 Oct. 2009.
- Lessig, Lawrence. *Free Culture*. New York: Penguin Group, 2003. Print.
- Masnack, Mike. "Focusing In On The Value: Google Books Provides An Amazing Resource." *Techdirt*. 2 Oct. 2009. Web. 3 Oct. 2009.
- Musgrove, Mike. "Google Settles Publishers' Lawsuit Over Book Offerings." *Washington Post*. 29 Oct. 2008. Web. 1 Oct. 2009.
- Perez, Juan. "In Google Book Settlement, Business Trumps Ideals." *PC World*. 30 Oct. 2008. Web. 1 Oct. 2009.
- "Report: Google to resume book scans." Editorial. *CNN Money*. 1 Nov. 2005. Web. 1 Oct. 2009.
- Rife, Martine Courant. "The fair use doctrine: History, application, and implications for (new media) writing teachers." *Computers and Composition*. 24.2 (2007): 154-178). Web.

Schwankert, Steven. "Microsoft hits Google over book search model." *MacWorld*. 6 March 2007. Web. 1 Oct. 2009.

*Static goods, dynamic bads*. Prod. Lawrence Lessig. Berkman Center, 2009. Web. 20. Nov. 2009. <http://www.youtube.com/watch?v=Svytkew5qPI>

U.S. Copyright Office. "Fair Use." *U.S. Copyright Office*. Web. 1. Oct. 2009.

Vaidhyanathan, Siva. "The Googlization of Everything." *Institute for the Future of the Book*. Web. 20 Nov. 2009. <http://www.googlizationofeverything.com/>

Yates, Carolyn. "Timeline of Google Book Search." *The McGill Tribune*. Web. 23 Nov. 2009.