Order Code RS22356
Updated January 22, 2007



CRS Report for Congress

# The Google Book Search Project: Is Online Indexing a Fair Use Under Copyright Law?

Robin Jeweler
Legislative Attorney
American Law Division

## Summary

Google, Inc. is digitally scanning the collections of several prominent libraries in order to create a vast searchable database of literary works. Copyright holders who have not authorized and object to the digitization have filed suit against the company. This report provides background on the pending litigation. It will be updated as judicial developments warrant.

Complaints for copyright infringement were filed against Google, Inc. by a variety of authors and representatives of the book publishing industry.[1] The complaints specifically challenge Google's "Print Library" project, renamed "Google Book Search," an effort by Google in conjunction with several library partners to scan books into a digital format so that they may be searched textually. Although the case is in its early stages, the issues presented have captured national attention. Once again, new technology and traditional principles of copyright law appear to be in conflict. Because of the unique facts and issues presented, there is scant legal precedent to legitimize Google's claim that its project is protected by copyright law's fair use exception to liability for infringement.[2] Thus, questions presented may be ones of first impression for the courts.

**Google's Book Search Project.** In December of 2004, Google announced a partnership with several major libraries to make digital copies of their collections and permit the text of the literature to be searched online by the Google search engine.[3] Google is providing its partnering libraries with a digital copy of the donor institution's collection.

---

[1] McGraw-Hill Companies, Inc. v. Google, No. 05 CV 8881 (S.D.N.Y., filed Oct. 19, 2005); The Author's Guild v. Google, No. 05 CV 8136 (S.D.N.Y., filed Sept. 20, 2005).

[2] 17 U.S.C. § 107.

[3] Participating libraries include those at the University of Michigan, Harvard University, Stanford University, Oxford, and the New York Public Library. For details, *see* [http://books.google.com/googlebooks/partners.html].

The Print Library Project is only one of several initiatives by the company to enhance the breadth of its online search capabilities. It was originally a component of "Google Book Search," which included the Library Project and its "Partner Program," an online book marketing program designed to help publishers and authors promote their books by displaying a limited number of sample pages in connection with a user's word search. The Library Project is described on its website:

> When you click on a search result for a book from the Library Project, you'll see basic bibliographic information about the book and, in many cases, a few snippets — a few sentences showing your search term in context. If the book is out of copyright, you'll be able to view and download the entire book. In all cases, you'll see links directing you to online bookstores where you can buy the book and libraries where you can borrow it.[4]

In addition, "Google Video" offers a search vehicle for material from archived television programs, educational videos, personal productions, and other video media. Users can search the closed captioning and text descriptions of its video archive for relevant results. However, in the case of both the Book Partner Program and the video program, content owners actively submit their material for inclusion in the searchable database. Only the Library Project does not seek authorization to copy from content owners. Hence, it is the only program that has been challenged in court.

After some academic and commercial publishers objected to the Library Project, Google took a brief hiatus from scanning to allow publishers time to identify works that they, i.e., the copyright holders, do not want to be included in the digital database. This has been referred to an "opt out" plan. The general rule of copyright law requires a prospective user to seek permission for use; Google has reversed the process by announcing its intention to digitize entire collections of the contributing libraries *unless* a content owner opts out by acting to *withhold* permission. This contributes to the content holders' claim that Google is engaged in massive copyright infringement.

**The Parties' Positions.** The complaint filed by plaintiff publishing companies (the Publishers) accuses defendant Google of massive copyright infringement. By digitizing copyrighted works without permission, Google is alleged to violate the copyright holders' exclusive rights to copy and/or display protected work.[5] Plaintiffs contend that Google's project is strictly commercial because it "pays" for the libraries' collections by delivering digital copies back to them; and, Google will realize significant advertising revenues as a consequence of its enhanced search capabilities.

Defendant Google essentially contends that its opt out program negates any infringement liability. But, if infringement were found, Google argues that its activity is protected by copyright's fair use doctrine. Google cites the U.S. Court of Appeals for the

---

[4] [http://books.google.com/googlebooks/library.html]. For more background on the program, see Google's website at [http://books.google.com/googlebooks/about.html].

[5] 17 U.S.C. § 106.

Ninth Circuit's decision in *Kelly v. Arriba Soft Corp.* as support for the proposition that Internet search engines' indexing activities constitute a fair use.[6]

**The Law.** As stated above, copyright law confers on the rights holder the exclusive right to control reproduction, display, and distribution of a protected work. Accordingly, in order to use a copyrighted work, one seeks permission from the owner and negotiates the terms, conditions, and payments for use. Google claims that for it to actively seek permission from every rights' holder in the multi-library collections would be impractical and prohibitive. The Publishers claim that Google's "opt out" program "stands copyright law on its head."[7] One cannot, they argue, generally announce one's intention to infringe multiple copyrighted works and collectively offer rights holders the opportunity *not* to have their work infringed. Yet, the requirement that a copyright owner act affirmatively to stop non-willful infringement is not without precedent. The DMCA's "notice and takedown" procedures — requiring the content owner to notify an Internet service provider (ISP) of the existence of infringing content — may immunize an ISP from infringement liability when it serves as a "passive conduit."[8] And at least one court has found that a content owner was responsible for taking affirmative measures, i.e, using meta-tags, to prevent Internet search engines from automatically searching and displaying the owner's Internet content.[9]

***Fair Use.*** Assuming a court were to find that Google's digitization of copyrighted works is infringing, the question becomes whether its activities are a fair use. The fair use exemption derives from common law and the First Amendment.[10] As codified in the Copyright Act, it establishes criteria for a court to consider in determining whether an infringing use is "fair." Specifically, 17 U.S.C. § 107 provides that, notwithstanding other provisions of the law, use of a copyrighted work

> ... for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include —
> (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.

---

[6] 336 F.3d 811 (9th Cir. 2003). For more background on Kelly v. Arriba Soft Corp., see CRS Report RL33810, *Internet Search Engines: Copyright's "Fair Use" in Reproduction and Public Display Rights* by Robin Jeweler.

[7] *Publishers: Value of Book Search Project Shows That Scanning Is Not Fair Use,* 71 BNA PATENT, TRADEMARK & COPYRIGHT J. 94 (Nov. 25. 2005).

[8] 11 U.S.C. § 512(b)-(c).

[9] *See* Field v. Google, 412 F. Supp.2d 1106 (D.Nev. 2006).

[10] Eldred v. Ashcroft, 537 U.S. 186, 219 (2003).

Because fair use is an "equitable rule of reason" to be applied in light of the overall purposes of the Copyright Act, other relevant factors may also be considered.[11]

The court hearing the case will make findings of fact and assign relative value and weight to the fair use factors in its analysis. This report will not attempt to predict an outcome in the pending litigation but does make some observations with respect to fair use analysis and the issues at hand.

The Library Project has the potential to be a great boon to scholarship, research, and the public in general. It is, nevertheless, commercial in nature because Google anticipates that it will enhance its service's utilization by the public and concomitantly increase advertising fees. With respect to the first factor, the purpose and character of use, the searching and indexing goal appears to be a highly transformative use of the copied text. There is little question that indexing basic information about any book alone, absent copying, would not constitute copyright infringement. While displaying "snippets" of text is closer to infringing activity, the prospective display, as described by Google, does not appear to usurp or negate the value of the underlying work.

The second factor is the nature of the copyrighted work. Digitizing the collections of the named libraries will encompass both factual and creative works, the latter being entitled to the highest level of copyright protection. How the court views the third factor — amount of the portion used — will be significant. In order to create its mega-database, Google will scan the entire copyrighted work, a major consideration weighing against fair use. But it intends to display, i.e., use, at any given time, only brief excerpts of the searchable text. Hence, is the digital reproduction incidental to an otherwise fair use or is it impermissibly infringing?

Finally, what will be the Library Project's effect on the potential market for or value of the copyrighted works? Here, Google makes a strong argument that its indexing and text searching capability has the potential to greatly enhance the market for sales for books that might otherwise be relegated to obscurity. Its "sampling" of text permits members of the public to determine whether they wish to acquire the book.

The Publishers counter that copyright owners routinely receive license fees for authorized sampling. Google's project may deprive them of the opportunity to participate in the creation of similar databases over which they have control and input. The Publishers have also expressed concern that the digital edition of the work Google returns to the participating library may facilitate piracy and/or additional unauthorized uses.[12]

***Case Law.*** Although a court's finding that there is a fair use exception to copyright infringement is context-specific, it naturally looks to precedent for guidance. Google asserts that *Kelly v. Arriba Soft Corp.* supports its claim of fair use, and in many respects it does. In *Kelly*, the court found that Arriba Soft's search engine, which, in response to a user's inquiry, compiled a database of images by copying pictures from websites ( and

---

[11] Sony Corp. of America v. Universal City Studios, 464 U.S. 417, 448, 454 (1984).

[12] In addition to fair use, the Copyright Act contains additional limitations on exclusive rights of copyright holders. 17 U.S.C. § 108, for example, gives certain reproduction rights to libraries and archives, but these provisions are not relevant to the issue discussed.

displayed them as thumbnail images, was a sufficiently transformative use of copyrighted material to be a fair use. Although providing indexing information alone does not implicate copyright infringement, displaying limited quotes from a literary work may be consistent with fair use. Indeed, a quote from a literary work is a far more limited reproduction and display than a thumbnail image of a full-sized one. However, a major distinction between the *Kelly* case and Google's Library Project is that in the former, content owners voluntarily uploaded their images to the Internet.[13] Here, the search engine is copying material to create a new database to enable the search capability. Whether a new database and "snippet" sampling will erode a viable new market for authors and publishers remains to be seen.[14]

Hence, the arguably unique question presented is whether apparent *prima facie* infringing activity that facilitates an arguably legitimate use is indeed a fair one. In this broad respect, the Supreme Court's decision in *Sony Corporation of America v. Universal City Studios*[15] is apposite. In *Sony*, the Court held that the sale of the video recording machine to "time shift" broadcast television for personal home viewing was not contributory copyright infringement. The Court articulated a new category of fair use, namely, time shifting. Although the factual underpinnings and legal precedent of this decision are *not* particularly relevant to nor controlling in the instant case, the *Sony* decision itself stands as a landmark in copyright law demonstrating the willingness of the Court to balance new technological capabilities against traditional principles of copyright law.

Although *Sony* sanctioned "time shifting" of in-home television broadcasting, neither the U.S. Supreme Court nor the lower courts have evidenced willingness to expand this judicially created category of fair use. In *UMG Recordings v. MP3.Com, Inc.,* a U.S. district court rejected out-of-hand the defendant's proffered fair use defense as a justification for unauthorized copying of plaintiffs' audio CDs.[16] The defendant claimed that its unauthorized copying enabled it to provide a service to proven owners of a CD, namely "space shifting," or enabling them to obtain access to the music on the CD through its subscription service.

In the case of Google, many copyright experts see analogies to the technological considerations inherent in *Sony*. The digital scanning — the alleged infringing activity — is viewed as incidental to a valid and socially useful function, indexing.[17] Courts

---

[13] Some observers suggest that uploading to the Internet may be construed as a limited license for foreseeable uses such as those by Internet search engines. Jonathan Band, *The Google Print Library Project: Fair or Foul?,* 9 J. OF INTERNET LAW 1, 4 (Oct. 2005). *See* Field v. Google, *supra* note 9.

[14] *See* Perfect 10 v. Google, 416 F. Supp. 2d 828 (C.D.Ca. 2006), holding that a search engine's display of thumbnail images may be infringing with respect to a content owner's developing market for the sale of reduced-size images.

[15] 464 U.S. 417 (1984).

[16] 92 F.Supp.2d 349, 352 (S.D.N.Y. 2000)("[D]efendant's 'fair use' defense is indefensible and must be denied as a matter of law.").

[17] *See, e.g.,* Jonathan Band, *supra;* Christopher Heun, *Courts Unlikely to Stop Google Book* (continued...)

appear increasingly willing to acknowledge the utility of online indexing as a component of Internet functionality.[18] Others view Google's activity as *prima facie* copyright infringement, with little or no extenuating circumstances.[19]

***Conclusion.*** How the court (or courts) that consider this case define the issues presented will ultimately determine whether the suit against Google sets an important precedent in copyright law. Viewed expansively, the court may find that copying to promote online searching and indexing of literary works is a fair use. To many observers, such a holding could be the jurisprudential equivalent of *Sony's* sanctioning of "time shifting." If the court adopts a more narrow view of fair use that precludes Google's digitization project, searchable literary databases are likely to evolve in a less comprehensive manner but with the input and control of rights holders who view them as desirable and participate accordingly.

GalleryWatch.com
http://www.crsdocuments.com
Penny Hill Press



---

[17] (...continued)
*Copying,* INTERNETWEEK (Sept. 2, 2005) at [http://internetweek.cmp.com/showArticle.jhtml?articleID=170700329].

[18] *See Field* and *Perfect 10, supra*.

[19] *See, e.g.,* Elisabeth Hanratty, *Google Library: Beyond Fair Use?,* 2005 DUKE L. & TECH. REV. 10 (2005); Raymond Nimmer, *Google Lawsuit Begins; Fair Use* (Oct. 3, 2005), online at [http://www.ipinfoblog.com/archives/intellectual-property-33-google-lawsuit-begins-fair-use.html].