The Google Book Search Project Litigation

"Massive Copyright Infringement" or "Fair Use"?

By Lawrence Jordan

Fast Fact:

Current litigation over Google's innovative Book Search Project raises interesting "fair use" and other legal issues.

Googling "Google" with "University of Michigan litigation" produces over 101,000 hits.[1] These involve the latest application of our aging Copyright Act[2] to new digital technology—in this case, Google's Library Project (the Project).

The legal issues in *Authors Guild v Google Inc*[3] (*Authors Guild Suit*) and *McGraw-Hill Cos, Inc v Google Inc*[4] (*Publishers' Suit*) include especially the application of the copyright doctrines of fair use and licensing to Google's ambitious book scanning and searching project.

The legal issues in *Authors Guild v Google Inc*[3] (*Authors Guild Suit*) and *McGraw-Hill Cos, Inc v Google Inc*[4] (*Publishers' Suit*) include especially the application of the copyright doctrines of fair use and licensing to Google's ambitious book scanning and searching project.

As with previous attempts to apply the 1976 Copyright Act to electronic copying technology, the courts are again being asked to apply a Copyright Act written for 1960s and 1970s technology to newer machines. The outcome is likely to test the flexibility and resilience of our Copyright Act and may support calls for amendment of the Act to expressly address some of the challenges raised by digital technology.

The Project

The Project plans to scan the text of millions of books and materials into a searchable database. The texts in question are in the collections of major U.S. and English libraries, notably the libraries of the University of Michigan, Harvard University, Stanford University, the University of Oxford, and the New York Public Library.[5] At the University of Michigan, Google plans to scan into a database works that are currently covered by copyright, as well as works that are in the public domain (generally, works published in the United States before 1912 are now in the public domain, as are certain other works). Google also plans to provide a copy of the database to each library.

The Project is expected to take 10 years and involve nothing less than creating searchable databases of most books in the English language. One estimate is that Google will scan 30 million books.[6] Understandably, a Google vice president has proudly called the Project a "man on the moon initiative."[7] Once the works are in the database, researchers will be able to search the content of the works. In addition, at least in theory, searchers (and hackers) would be able to copy entire texts.

Not all parties are as enthusiastic as Google about the Project. The *Authors Guild Suit,* for example, characterizes the Project as "massive copyright infringement."[8]

Partly in response to complaints from copyright holders, Google has developed a two-prong approach to handling copyrighted works: (1) limiting to a small "snippet" the amount of copyrightable text that appears to a searcher (a "snippet" being a small portion of text on either side of the searched term), and (2) allowing copyright holders to "opt-out" of the program. By contrast, public domain text can be viewed in its entirety.

Google initiated the "opt-out" program in August 2005. Under this program, a copyright holder could demand that its works be removed from the Project. In the New York litigation, the copyright holders complain that this approach turns copyright law on its head. They assert that they should not be required to affirmatively "opt-out." Rather, they assert, the burden should be on a potential infringer to respect the rights of copyright holders.

The New York Cases

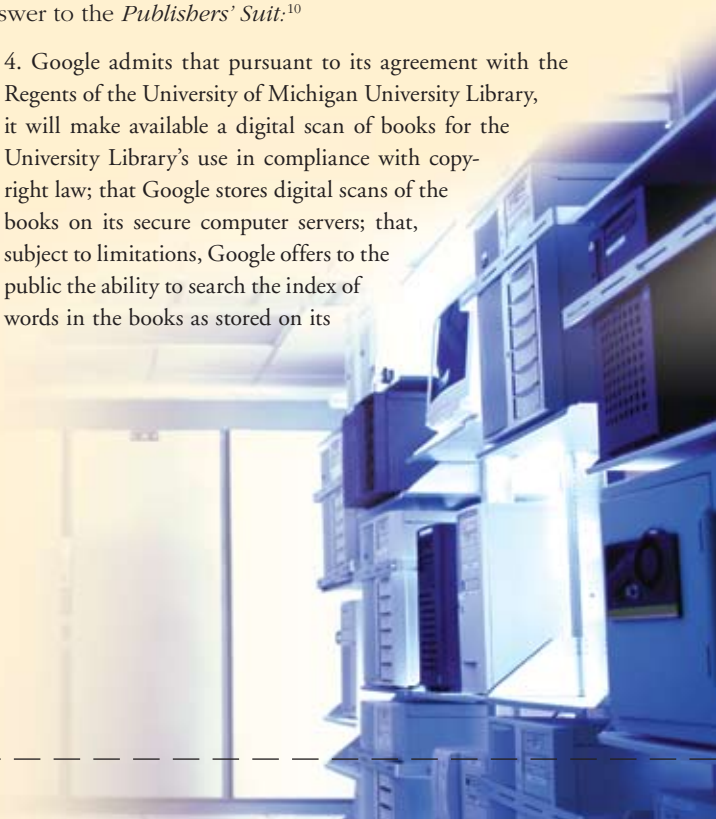
While the *Authors Guild Suit* and the *Publishers' Suit* differ in some respects, they both assert copyright infringement and seek permanent injunctive relief against Google. Discovery and other pre-trial procedures have been consolidated in front of Judge John E. Sprizzo pursuant to an order of coordination. Motions for summary judgment were to be filed by July 2, 2007. Trial is expected in 2008.

The position of the plaintiffs in both suits is, in essence, that Google is engaged in copyright infringement for the purpose of increasing its advertising revenue. As stated in the *Publishers' Suit* complaint:[9]

> 4. In consideration for receiving books from Michigan for scanning, Google proposes to make a digital copy of each book that it scans and then provide that copy to Michigan for Michigan's own use. Google also proposes to (a) store, in perpetuity, one or more of the resulting digital copies on Google's computer servers, (b) offer to the public the ability to search, and have access to, the copies of the books stored on Google's servers and to retrieve excerpts of those books and (c) publicly display the excerpts of the books to any person in the world whose search, through Google, has retrieved that book. All of these steps are taken by Google for the purpose of increasing the number of visitors to the google.com website and, in turn, Google's already substantial advertising revenue.

Google's position is largely summarized in paragraph 4 of its answer to the *Publishers' Suit:*[10]

> 4. Google admits that pursuant to its agreement with the Regents of the University of Michigan University Library, it will make available a digital scan of books for the University Library's use in compliance with copyright law; that Google stores digital scans of the books on its secure computer servers; that, subject to limitations, Google offers to the public the ability to search the index of words in the books as stored on its



secure servers; that a list of books matching an individual's search is returned to the individual upon searching; that the information provided to the individual conducting the search with respect to a particular book varies depending on numerous criteria, including the copyright status of the book and whether the book was submitted to Google for hosting under contract with the holder of its copyright; and that for books subject to copyright and not submitted by the copyright holder, Google will display (among other things) bibliographic information and, at most and depending on the type of book, three or fewer very short excerpts of approximately one vertical inch each.

The Fair Use Defense

Google's principal (though not only) substantive defense in both cases is "fair use."[11] This long-standing copyright doctrine allows limited copying of copywritten works, "for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research…." The statute requires that the court analyze a fair use defense using four factors: (1) "[t]he purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit education 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."

How Will the New York District Court Apply the Statute to the Present Litigation?

Two lines of cases suggest alternative outcomes. While there is no controlling Second Circuit case, the Southern District of New York ruled in *UMG Recordings v MP3.com*[12] that a defendant that copied copywritten works onto a server, and allowed access to third-party subscribers, was not protected by fair use. In the *UMG* case, defendant provided a "space-shifting" service, under which customers who already owned copies of music in CD form could store it on the UMG server. The court held that copying the copywritten content into the database was not a "fair use and, accordingly, constituted copyright infringement."

By contrast, defendants would argue that a Ninth Circuit case, *Kelly v Arriba Soft,*[13] is more on point. In *Kelly,* the defendant used a software spider to locate photographic images on the Internet. These images were reduced to "thumbnails" and stored on defendant's database, where they could be searched. In upholding the fair use defense, the Ninth Circuit held that Arriba's use of "thumbnails" was "transformative" and did more than merely copy the images. The court accordingly found that the first fair use factor favored the defendant. In addition, the *Kelly* court found that making it easier to find the copywritten works on the Internet did not diminish the value of the work (the fourth statutory factor). Defendant's position has been strengthened by the Ninth Circuit decision in *Perfect 10, Inc v Amazon.com, Inc.*[14] The *Perfect 10* court upheld Google's fair use defense, ruling that the district court erred in holding that use of thumbnails was copyright infringement.

Summary

No one knows how the Google litigation will be resolved. Perhaps it will produce a significant precedent in the law of fair use; perhaps a settlement with significant economic impact on the distributors of knowledge. In any case, the litigation has already earned a place in the evolution of copyright law. One wonders what databases we'll be able to use to access the eventual outcome. ■

Lawrence Jordan is a member of Jaffe, Ratt, Heuer, and Weiss, where he heads the Ann Arbor office and the Intellectual Property Practice Group. He is a past chairperson of the State Bar of Michigan Arts, Communications, Entertainment, and Sports Section and the Computer Law Section, and a member of The Institute of Continuing Legal Education Technology Advisory Board.

FOOTNOTES

1. <http://www.google.com/search?hl=en&q=%22+22University+of+Michigan%22+litigation8btn6=Google+Search> (accessed May 7, 2007).
2. 17 USC 101, et seq.
3. *Authors Guild v Google Inc,* United States District Court for the Southern District of New York, Docket No 2005 CV 8136, filed September 20, 2005.
4. *McGraw-Hill Cos, Inc v Google Inc,* United States District Court for the Southern District of New York, Docket No 2005 CV 08881, filed October 19, 2005. The other plaintiffs include Pearson Education, Inc.; Penguin Group (USA) Inc; Simon & Schuster, Inc.; and John Wiley & Sons, Inc.
5. On June 6, 2007, the Committee on Institutional Cooperation (CIC) announced an agreement with Google, which allows the digitization of the distinctive portions of the CIC members' library collection. The CIC includes 12 midwestern universities, including Michigan State University.
6. Jonathan Brand, *The Google Library Project: The Copyright Debate,* OITP Technology Policy Brief, American Library Association, January 2006.
7. Google Book Search: All the world's books at your fingertips <http://books.google.com/googlebooks/vision.html> (accessed August 21, 2007).
8. Footnote to ¶ 3 of the complaint in the *Authors Guild Suit.*
9. Complaint in *McGraw-Hill,* n 4, *supra.*
10. Answer in *McGraw-Hill,* n 4, *supra.*
11. 17 USC 107.
12. *UMG Recordings v MP3.com,* 92 F Supp 2d 349 (SD NY, 2000).
13. *Kelly v Arriba Soft,* 336 F 3d 811 (CA 9, 2003).
14. *Perfect 10, Inc v Amazon.com, Inc,* 487 F 3d 701 (CA 9, 2007).