Google Will Return to Scanning

Copyrighted Library Books

By Kevin J. Delaney
And Jeffrey A. Trachtenberg

GOOGLE INC. will resume scanning copyrighted library books into its search engine after a self-imposed hiatus, despite the efforts of some publishers and authors to block it from doing so without the copyright holders' permission.

The Mountain View, Calif., company said it plans to resume scanning copyrighted books in the collections of Stanford University and the University of Michigan "soon." In August, amid a backlash by publishers, the company suspended its scanning of copyrighted books in libraries until today in order to give publishers time to request that their works not be scanned.

Google also said it will focus on scanning copyrighted works that are out of print and is seeking publishers' permission to digitize books that are still available new from bookstores. The company hadn't publicly disclosed that it was targeting out-of-print works for its initial scanning efforts although it says that was always its policy. That distinction could strengthen its argument that the project won't negatively affect book sales. But it isn't likely to lead the two groups suing Google over the Print Library Project to back down.

Google's renewed push is the latest move in a legal and philosophical battle over how copyright law applies to online content, a battle that could have ramifications for the future of the Internet and the publishing industry. Through its Print Library Project, which was announced in December, Google plans to put millions of library books through digital scanners that record images of each page. It will use those scans to let consumers search for keyword phrases within the text of books and see relevant portions of the text. The program is separate from the Google Print Publisher Program, under which publishers give permission for Google to make their books available through its search engine, providing Google with scans or books for it to scan. With that program, publishers can request that Google display only bibliographic information or a small excerpt of the text to users.

Google's critics have objected to the company's plan to scan the library books without the authors' or publishers' permission, requiring copyright holders to proactively "opt out" of the scanning. At its core, the two sides are scrapping over the online application of "fair use," a legal doctrine allowing the use of copyrighted material for certain purposes, including teaching, research and news reporting. Some lawyers believe at least one of the lawsuits against Google could make its way through the legal system to the higher courts, eventually setting a precedent for fair use in the Internet age.

"We prefer to work with the publisher for in-copyright material," said Susan Wojcicki, a Google vice president for product management. She said that Google is first scanning books from library shelves that primarily contain out-of-print works, though some in-print books on such shelves could be included inadvertently. Because Google wants to be able to search all books, Ms. Wojcicki said it will eventually scan in-print material even when it doesn't have the author's or publisher's consent, though Google "would never show a full page without the right from a copyright holder."

Publishers and authors suing the company suggest that Google's focus on scanning out-of-print works isn't likely to soften their legal opposition to it. Paul Aiken, executive director of the Authors Guild, said his trade group doesn't intend to withdraw the suit it filed in September against Google in U.S. District Court in New York alleging that the search company is violating copyright.

Please Turn to Page B5, Column 1

Continued From Page B1

"Books that are out of print frequently come back in print," said Mr. Aiken. "A university press or a smaller house may bring it back, or it may come back when the author publishes a new book with a major publisher."

"The fact that we filed suit shows how important the issue is to us," said Allan Adler, a lawyer at the Association of American Publishers, a trade group that last month filed suit on behalf of several members against Google, also in federal court in New York. "It may be unwise to make any significant change in our posture."

He noted that the issues surrounding books that are out of print but in copyright remain sticky. "This category is fraught with issues as to whether publishers can even say go ahead and use them because they may not have the rights," said Mr. Adler.

It isn't known just what percentage of library holdings fall into the category of being in copyright but out of print. About 18% of the books held by the libraries working with Google were printed prior to 1923 and are therefore in the public domain, according to an analysis by the Online Computer Library Center, a Dublin, Ohio, nonprofit library cooperative. An unknown percentage of the rest still are protected by copyright, depending on whether it was renewed.

Google's resumption of its scanning of copyrighted works comes amid heated debate in the library community over participation in the program. "We believe the Google project is a good and right thing, and that our participation is consistent with our role as an educational institution," said Andrew Herkovic, director of communications and development at the Stanford...
University Libraries. Stanford has provided Google with volumes that are in the public domain, but Mr. Herkovic declined to say whether the university has also provided copyrighted material to Google.

Other librarians say that what matters most is that people have as much access to information as possible. "We think that what Google is doing is legal and consistent with copyright law because copyright law is about striking a balance between the limited rights of the copyright owner and the long-term rights of the public," said John Wilkin, associate university librarian at the University of Michigan, which is making an estimated seven million volumes available.

But Google's actions have raised concerns among some librarians uneasy with the idea of creating ill will with publishers and authors. "In general I think that libraries need to observe copyright," said Tom Garnett, assistant director for Digital Library and Information Systems at the Smithsonian Institution Libraries, which are the libraries serving the Smithsonian Institution. Mr. Garnett said that the Smithsonian Institution Libraries will maintain a conservative stance until the disputes are resolved.

Other librarians are more strongly opposed to the Google project. "I feel that this is a potential disaster on several levels," said Michael Gorman, president of the American Library Association and university librarian at California State University, Fresno. "They are reducing scholarly texts to paragraphs. The point of a scholarly text is they are written to be read sequentially from beginning to end, making an argument and engaging you in dialogue."

Mr. Gorman, who said the American Library Association doesn't have an official position on the subject, described Google's argument that Web users will be able to look at several snippets and then decide whether they want to buy or read the book as "ridiculous." Further, he noted that as a published author, he opposes Google's intention to build an enormous database that includes copyrighted texts. "It's a flaunt of my intellectual property rights," he said.