



Will the Online Book Publishing Flap Rewrite Copyright Law?

Published : February 08, 2006 in [Knowledge@Wharton](#)

The latest frontier in the digital content revolution -- efforts by Google, Amazon and others to turn millions of books into bytes that can be accessed no matter where the reader is, sold by the page and easily searched -- could redefine copyright law and change the way knowledge is shared around the world, say experts at Wharton.

After Google announced Google Print -- its plan to digitize the libraries of Harvard, the University of Michigan, Stanford, Oxford and the New York Public Library -- the Authors Guild and the American Association of Publishers filed lawsuits over the scanning of copyrighted books. Their action could lead to a high profile court case redefining copyright law, which now states that a content owner's permission is required before intellectual property can be copied, redistributed, displayed or used to form the basis for derivative works.



This is a single/personal use copy of Knowledge@Wharton. For multiple copies, custom reprints, e-prints, posters or plaques, please contact PARS International: reprints@parsintl.com P. (212) 221-9595 x407.

On November 3, Google launched the first installment of Google Print, which dodged the copyright issue by including only texts which were in the public domain. Because this material is no longer protected by copyright law, anyone can browse the full contents of a book and save individual pages. The project's launch, however, was overshadowed by news that the company plans to resume scanning copyrighted works to expand the Google Print database. Google had originally begun doing this in August, but delayed the process to allow publishers the option to exclude their works. Google plans to put copyrighted books in its search results, but limit what users will see unless it gets permission from publishers and authors to show more. Publishers can choose to keep their books digitized by Google and its search engine.

"Google is clearly going out on a limb with respect to copyright. The limb may well hold, which I think would be the better result as a matter of public policy," says Wharton legal studies and business ethics professor [Kevin Werbach](#). "On the other hand, the limb could easily break. The courts will decide."

Indeed, Google's recent move to scan copyrighted works leaves many unresolved questions: Does the greater good of putting books online outweigh current copyright law? Is Google's complete scanning a violation of copyright law even if the end user doesn't get much more than a small excerpt of the work in a search result? Does it make a difference if the book is out of print yet still, theoretically, under copyright? Should Google be required to get a publisher's permission before scanning content rather than offering an opt-out policy that puts the burden on the publishers to take action? Is copyright law designed for printed materials still valid in the digital age?

No matter how those copyright issues are resolved, Google will hardly be the only player in this arena. Just days after Google Print launched, Amazon introduced two new programs -- Amazon Pages and Amazon Upgrade. With its recent move, Amazon hopes to do for books what Apple Computer's iTunes did for music: Allow a consumer to buy an entire book, or just parts of it, much like consumers can now buy a specific song without buying the entire album. Amazon Pages allows a buyer to purchase a chapter or even a relevant page. Amazon Upgrade enables online access to any book that is purchased.

Random House has announced that it will work with online booksellers like Amazon, search engines like

Google and portals like Yahoo to offer contents of its books "a la carte" for small payments, such as 99 cents for four pages. Participants as a general rule will share revenue, although specific pricing plans will depend on negotiations with partners, said Random House. "We believe that it is important for publishers to be innovative in providing digital options for consumers to access our content," stated Richard Sarnoff, president of the Random House corporate development group, in a statement.

Meanwhile, *The Wall Street Journal* reported on November 14 that Google is talking to publishers to gauge interest in a potential plan to rent digital versions of books to consumers.

And Microsoft reportedly plans to digitize the British Library (the national library of the United Kingdom), while Yahoo is leading a group called the Open Content Alliance whose goal is to create a digital archive of "globally sourced digital collections, including multimedia content." The Open Content Alliance, however, won't include works in its archive without permission.

According to Wharton marketing professor [Peter Fader](#), it's hard to compare the respective plans of Google and Amazon. For instance, Amazon's effort is more utilitarian: You pay for a page you need. Google's effort is more about disseminating as much information as possible. "The two approaches are entirely different," he says. "There is a time and place for each one."

Werbach points to Google's "breathtaking ambition to organize the world's information. Scanning books into its database is the next logical step toward that goal. Once Google made a big commitment, its competitors had to jump in as well. Books are a truly vast source of knowledge that isn't nearly as accessible as the materials on the web."

From Snippets to Chapters

The battle of digitizing books isn't that far removed from other skirmishes between old distribution models and new ones for music and video. On the music front, Apple Computer has carved out a dominant market position by being a legal repository for digital music downloads. Disney's ABC network recently announced a deal to distribute some of its shows, including "Desperate Housewives," over iPods. CBS and NBC followed up with plans to offer shows on demand for 99 cents on Comcast and DirecTV, respectively.

But unlike Apple, which got music publishers to buy into its digital distribution scheme before launching iTunes, Google seems poised to charge ahead without prior approval from the copyright holders. Consequently, on October 19, the Association of American Publishers announced that it was suing the company, a suit organized and funded by the AAP and joined by The McGraw-Hill Companies, Pearson Education, Penguin Group (USA), Simon & Schuster and John Wiley & Sons. According to AAP president and former Colorado Congresswoman Patricia Schroeder, the biggest worry is that Google's scanning of copyrighted works will water down intellectual property protection.

Google has answered critics on its corporate blog. In response to a suit filed by the Author's Guild on September 20, Susan Wojcicki, Google vice president, product management, says the company doesn't show "even a single page to users who find copyrighted books through this program." Unless a copyright holder gives permission, Google shows "a brief snippet of text" where a search term appears, along with basic bibliographic information and several links to online booksellers and libraries. Wojcicki adds that "Google respects copyright" and she likens Google Print to an electronic card catalog.

Schroeder, in an interview with Knowledge@Wharton, responds that "snippet" isn't a legal term and could evolve from meaning a sentence to meaning a complete chapter. Google's argument that publishers can opt out rings hollow, she adds. "The law doesn't say [an author or publisher] can opt out. It says you have to get permission before copying." Schroeder argues that if an opt-out policy became the norm, the burden would be on publishers to monitor usage and protect copyright, a burden she says would be

unmanageable.

According to R. Polk Wagner, a law professor at the University of Pennsylvania, the case between publishers and Google is a legal toss up. For instance, Google won't sell the book directly, but will point to places where the book can be purchased. Google's indexing could also provide some economic benefit to a copyright holder, especially if a book browser becomes a book buyer. In addition, Google can argue that its indexing of books is fair use and better serves the public good. One argument that will be made on Google's side is that its digital library effort serves the public's need for information, notes Werbach.

Wagner, however, suggests that the book industry is wary of Google hosting copyrighted works on its servers and becoming the de facto distributor of publishers' content. Google, according to Schroeder, should get permission before building what could be a profitable business around being a gatekeeper to books.

The fracas isn't likely to be resolved soon. "Either way, the case may well go all the way to the Supreme Court, which means it will be at least two years before we know the answer," says Werbach.

Seven Million Volumes in Seven Years

Amazon's approach is more in line with traditional notions of intellectual property, says [Dan Hunter](#), legal studies and business ethics professor at Wharton. By developing a system in cooperation with publishers, the company is staying above the fray and could boost its sales. "Amazon has relationships with publishers and that means that their search-inside-the-book function has attracted less ire than Google," says Hunter. "Google seems to be taking the approach that this is good for society and for them, so they will risk it even with the ire of the authors. Amazon has a much closer relationship with the publishers so it appears to be in a less fractious situation."

According to [Eric Clemons](#), Wharton professor of operations and information management, Amazon may be angling for a niche market of people who will want just a page or two of a book. "Amazon is targeting an audience that only wants to buy a chapter and not an entire book," says Clemons. "I think that this is a tiny market." Nevertheless, the business model may work just like Apple's iTunes did. "iTunes is priced principally to sell iPods, not to make much money off music sales. If Amazon uses this to get me to decide to buy the whole book, and gives me credit for the online chapter I bought as a teaser, they might increase book sales."

Another possible factor in the book industry equation is the power of libraries, some of which have embraced Google's digitization methods. The goal of libraries is to make their collections open to all. Digitization allows libraries to have a larger audience even if there are copyright challenges in the future. Indeed, Mary Sue Coleman, president of the University of Michigan, has said she welcomes the world to the school's library. "As educators, we are inspired by the possibility of sharing these important works with people around the globe," she noted in Google's November 3 statement. "Think of the doors it will open for students; geographical distance will no longer hamper research."

The flurry of activity around digitizing books is a bit surprising, say Wharton experts. The most obvious question is why weren't these collections online already. John Mark Ockerbloom, digital library architect and planner for the University of Pennsylvania Library, says the need to scan books and put them online isn't exactly a new development. "We have known for a while that users were interested in getting our information in digital form. The problem has been developing the workflow and digitizing lots of material quickly."

Indeed, the University of Michigan notes on its web site that its deal with Google will enable it to digitize its collection of seven million volumes in seven years. Under the arrangement, Google will scan a copy for its database and another for the university. Michigan launched an effort in 1995 to scan its books and

put them online, but so far has completed only 9,000 books. At that pace, the university says, it would have taken "more than 1,000 years to digitize our current collections."

While it's unclear what technology Google has developed to digitize books at such a fast clip, Ockerbloom says rapid high volume scanning that doesn't damage the books in a collection is a big selling point for the libraries. "This is huge," he notes. "Previously, the scanning required a lot of page flipping by hand. An effort like this will be closely watched." Digitizing books on the shelf "is an enormous task," adds Polk. "Search and storage technology only has recently evolved to a point where it is capable of handling the influx of data. Law firms have been dealing with this for a while and it is enormously expensive. Just digitizing the books published [from today on] is a [major] undertaking."

What's at stake for book publishers could be the economics underpinning the industry for the last 150 years, says [Daniel Raff](#), a Wharton management professor. The book industry depends on producing books, building inventory and then selling it. In between, wholesalers, retailers and publishing houses take their cuts. If digital delivery -- through Google, Amazon or otherwise -- becomes common, the industry could move from producing books ahead of demand to making them on demand, Raff notes. "The real bottom line is that the book industry has been around since the earliest days of the republic and now it's at a moment in history where technology is forcing dramatic change. It's evolution playing out in slow motion."

Wagner says it will take time to sort out the copyright and business issues surrounding the digitization of books. In his opinion, the best outcome of the Google-publishing industry dispute is that the court case that has been filed will establish some copyright ground rules in the digital world. "We will get to see how copyright law works in the real world."

Nevertheless, any clarity will likely be short-lived. "We always think there will be this clear moment, but it's not true in this context," Wagner adds. "It's an uncertain time and as soon as you answer the Google question [about copyright law] there will be another."

This is a single/personal use copy of Knowledge@Wharton. For multiple copies, custom reprints, e-prints, posters or plaques, please contact PARS International: reprints@parsintl.com P. (212) 221-9595 x407.