



DOI:10.1145/2566965

Pamela Samuelson

## Legally Speaking Mass Digitization as Fair Use

*Considering the implications of the late-2013 ruling in favor of Google in the Authors Guild case.*

**G**OOGLE HAS SCANNED 20 million books for its Google Book Search (GBS) project. In most countries of the world, this mass digitization would be deemed copyright infringement as to books not in the public domain. In the U.S., however, a doctrine known as fair use makes it possible to argue that scanning books to make an index and to provide snippets of their contents in response to search queries does not infringe copyrights.

The Authors Guild, a nonprofit organization representing approximately 8,500 professional authors, believes that Google's scanning of in-copyright books from the collections of research library partners is copyright infringement. In 2005, it brought a lawsuit to challenge this practice.

In 2008, Google announced a settlement of this lawsuit as well as a similar lawsuit brought by five trade publishers. That settlement would have given Google the right, among other things, to commercialize all out-of-print books in the GBS corpus through

three different business models and to display up to 20% of book contents in response to search queries unless rights holders affirmatively stepped forward to opt out of these business models and displays. Judge Denny Chin disapproved the proposed settlement in March 2011.

After further negotiations failed to produce a new settlement, Judge Chin established a schedule for proceeding with the litigation. Accordingly, the Authors Guild and Google filed motions for summary judgment. (This allows judges to resolve disputes if no key facts are in dispute and the litigants just disagree about how the law applies to those facts).

Less than two months after oral argument on these motions, Judge Chin ruled in November 2013 that Google's scanning, indexing, and snippet-providing were fair and non-infringing uses of the in-copyright books in the GBS corpus. The Authors Guild has announced it will appeal. This column will explain why I predict the appellate court review will result in an affirmation of Judge Chin's decision.

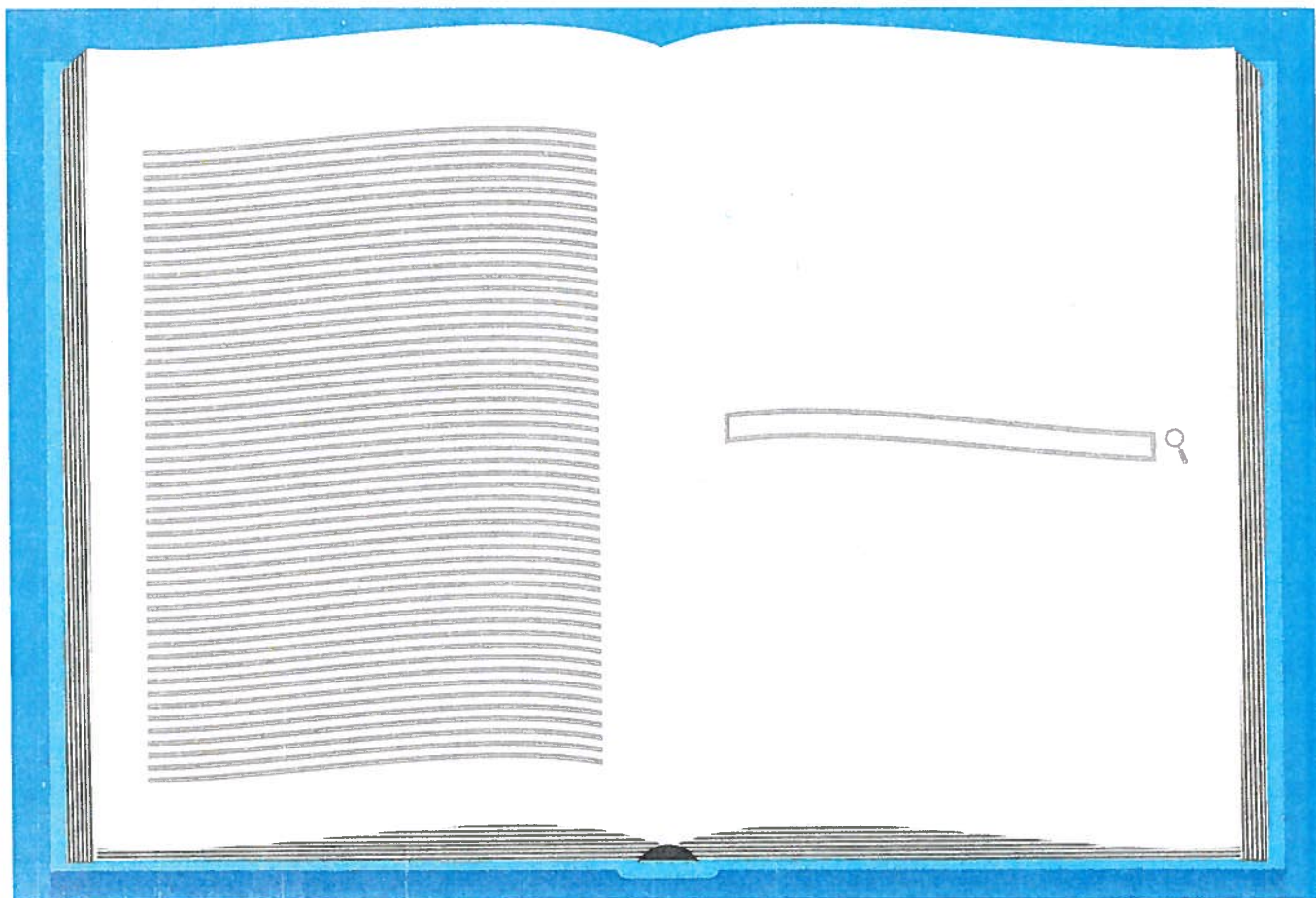
### What Is Fair Use?

Making fair uses of copyrighted works is lawful under U.S. law. Courts typically consider four factors in deciding whether a challenged use is fair or infringing: the purpose of the defendant's use; the nature of the copyrighted work; the amount and substantiality of the taking; and the harm to the market for the work.

Quoting a few sentences for a book review, parodying a popular song, reproducing parts of a politician's speech for news coverage, photocopying a news article relevant to a research project, and making time-shift copies of television programs are among the conventional uses that U.S. law would generally consider fair.

Most countries in the world do not have a flexible doctrine such as fair use in their copyright laws. Instead, legislatures typically identify specific types of uses that should be exempted. Quotation, news reporting, parody, and private study privileges are common in national copyright laws.

What is different about fair use is that it can be—and often is—applied



to situations not contemplated by the legislature. When Congress enacted a new copyright law in 1976, it did not think about whether making time-shift copies of television programs should be treated as infringement, let alone whether makers of videotape recorders (VTRs) should be held indirectly liable for any infringing uses of the devices. Courts were able to look to fair use as a way to balance the interests of copyright owners and users of VTRs and reach a judgment on these issues.

The Supreme Court's 1984 *Sony v. Universal City Studios* decision ruled that time-shift copying was fair use. The use was private and noncommercial. The programs copied had been shown on broadcast television for free. Whole programs were typically copied but consumers usually taped over the copy after seeing the program at a later time so the copies were not being retained. Universal had conceded no harm had yet occurred from time-shifting. Evidence offered about possible future harm was, in the Court's view, too speculative to matter.

Because Sony's VTR had substantial non-infringing uses, the Court concluded that consumers should be able to buy VTRs for these uses. Hence, Sony was not liable for infringing uses by some VTR users.

Subsequent fair use cases have addressed similar challenges posed by technological advances, including MP3 players, add-on software, and search engines.

### The Authors Guild's Position

Although Google is not directly monetizing the books whose contents it displays when serving up snippets from in-copyright books, the Guild has asserted Google had unquestionably scanned the books for commercial purposes (for example, to improve its search engine technologies). Commercial purposes usually cut against fair use.

Unlike the reviewer who quotes from a book or a musician who parodies a popular song, Google is not creating a new work of authorship that builds upon pre-existing works and contributes to knowledge or culture.

Because of this, the Guild argues that Google's use of the books is "non-transformative." This also tends to cut against fair use.

Google has, moreover, made copies of millions of highly creative books without seeking rights holder permissions. Indeed, it has made multiple copies of each book in the corpus. And it has archived these copies on its servers. All of these factors, in the Guild's view, disfavor fair use.

The Authors Guild has raised three main arguments about harm. First, it invoked Supreme Court decisions saying that harm should be presumed when defendants make non-transformative uses of protected works for commercial purposes.

Second, the Guild has asserted that Google is interfering with a market that would have developed to license the kinds of uses that Google has made of these books. The Guild has contended that this potential licensing market has been harmed.

Third, the Guild has pointed to the risk that clever hackers might break into Google servers and "liberate" mil-

lions of in-copyright books, thereby harming the market for purchases of books. Or clever researchers could display much of a particular book's contents by running a series of search queries that would allow them to consume a substantial amount of the book without purchasing it.

In the Guild's view, no one should be able to make systematic copies of entire books without seeking copyright owner permission.

### Judge Chin Agreed With Google

While accepting that there was some commerciality in the GBS project, Judge Chin agreed with Google that it had made "transformative" uses of the books. The Supreme Court has interpreted this term as including uses for a different purpose than the original. Google "transforms [the] expressive text [of books] into a comprehensive word index that helps readers, scholars, researchers, and others [to] find books," so its use of the books was not just transformative, but highly so.

Judge Chin relied upon some prior decisions saying that search engine "thumbnails" of images found on the open Web were transformative because they improved access to information. This fulfills the U.S. constitutional purposes of copyright law ("to promote the progress of science and useful arts"). Snippets of text, like thumbnails, enable users to find works in which they are interested. Thus, the purpose-of-the-use factor weighed strongly in favor of fair use.

The nature-of-the-work factor was not given much weight, although Judge Chin noted the vast majority of GBS books were non-fiction books from research library collections that were out-of-print. This somewhat favored fair use.

The Guild's harm assertions did not persuade Judge Chin. It was more likely, in his view, that GBS would be beneficial to authors and publishers. After all, GBS enables users to discover books in which they are interested and provides links to websites of booksellers so they can purchase relevant books.

The one factor that disfavored fair use, albeit only slightly, was the amount-and-substantiality-of-the-taking factor. Whole books were unquestionably copied. Yet

## The Guild has asserted Google had unquestionably scanned the books for commercial purposes.

Judge Chin recognized this copying was necessary to make the index. It also mattered that Google displayed only a few snippets from each book in response to search queries.

### Secondary Liability?

The Authors Guild claimed that Google was secondarily liable for copyright infringement because it provided its research library partners with digital copies of books that Google had scanned from their collections. The libraries have pooled these copies into a digital repository known as the HathiTrust.

The Guild brought a separate copyright infringement lawsuit against HathiTrust and Google's library partners for uses of books in the HathiTrust corpus. HathiTrust, like Google, has principally relied on fair use to justify its actions.

HathiTrust has asserted three purposes as favoring its fair use defense. Digitization preserves books in partner library collections. It enables text-mining so that researchers can discover books relevant to their interests. It enhances access to books for print-disabled persons. Once books are in digital form, it is relatively simple to convert the texts into braille or aural forms.

The *Authors Guild v. HathiTrust* case, like the *Authors Guild v. Google* case, was decided by a trial court on a summary judgment motion. Judge Baer, who presided over the *HathiTrust* case, decided in October 2012 that HathiTrust's uses of the digital books were fair.


In considering the secondary liability claim in the Google case, Judge Chin agreed with Judge Baer that the HathiTrust uses were fair. Google cannot be held secondarily liable for copyright infringement if the libraries' uses are fair, as Judges Baer and Chin have decided they are.

### What Will Happen on Appeal?

The Authors Guild has appealed its loss in the *HathiTrust* case and plans to appeal in the *Google* case. A panel of three appellate court judges heard oral argument in the *HathiTrust* case in October 2013. Two of the judges seemed quite interested in and supportive of HathiTrust's fair use defense.

These two judges will also be on the panel that hears the Guild's appeal in the *Google* case. The third judge that will be on the *Google* panel is a famous fair use expert. He was sympathetic toward Google's fair use defense when that panel overturned a class certification ruling from which Google had appealed. This panel of judges ordered the fair use appeal to come back to them. One never knows, of course, what will actually happen on any appeal, but my prediction is the appellate judges will affirm both fair use rulings.

This is not only good news for Google, but also for U.S.-based libraries, archives, historical societies, and similar institutions that want to undertake mass-digitization projects to increase public access to their collections.

Even though these institutions might be confident that mass-digitization projects would, if tested in litigation, be deemed fair uses, they typically lack financial resources to defend fair uses. They may also be intimidated by the risk of large statutory damage awards if they lost the cases. Google has both the resources and will to take on this monumental struggle and win. Of course, Google did not fight the Authors Guild's lawsuit in order to benefit the public, but if this is a by-product of the ruling in the case, so much the better. 

Pamela Samuelson (pam@law.berkeley.edu) is the Richard M. Sherman Distinguished Professor of Law and Information at the University of California, Berkeley.

Copyright held by Author/Owner(s).