

# Cites & Insights

Crawford at Large

Libraries • Policy • Technology • Media

Sponsored by YBP Library Services

Volume 9, Number 4: March 2009

ISSN 1534-0937

Walt Crawford

## Perspective: The Google Books Search Settlement

Once upon a time (late 2004), a young upstart company around these parts began an ambitious project to scan (digitize) and index old library books. Google was the name; the Google Library Project was the game. Several big important libraries signed up, some to scan everything, some only books in the public domain, some only small pilot projects. As time went on, more libraries signed on.

Why not? Google covered the costs and did the hard work and the libraries received copies of the scans. If it all worked out, the books would be much more discoverable through Google, making the library collections more useful.

Reduced to its essentials, the Google Library Project would yield two public faces as scanning went on and more books became findable through full-text words and phrases:

- For books in the public domain, fully readable and downloadable copies (although the quality of scanning was such that “fully readable” was sometimes more promise than reality).
- For all other books, “snippets” showing just a line or two of text—coupled with information on ways to get to the actual books.

These complemented Google’s existing Publisher Project, which made many in-print books findable and showed a few pages of the books (if the publishers said it was OK).

But some publishers and some authors were unhappy. Google hadn’t consulted them on this project. So some publishers (Association of American Publishers) and some authors (Authors Guild) filed a class-action lawsuit in 2005, claiming Google was infringing copyright. Google claimed that the scanning and snippets represented fair use. Scholars, pundits and other folk (like me) came down on all sides of the issue. The suit was in court for a long time. I wrote about the issues as part of ongoing coverage of Google

Book Search (and complementary or competitive projects, e.g. the Open Content Alliance).

But you know all this already, right?

Seven million books and a few years later, in late October 2008, Google announced a proposed settlement of the lawsuit. While the settlement hasn’t received final approval, it’s on its way. That’s the genesis for this PERSPECTIVE, written a few months after the announcement—long enough for most serious pundits to weigh in on the issues and the outcomes.

### My First Take

I *deliberately* avoided reading the commentaries or the settlement in any great detail. I knew it was premature to do any kind of overview and didn’t want to prejudge the situation. Instead, I glimpsed at the first few paragraphs of some four dozen commentaries and printed off the usual first page so I could come back to them later. That stack of first pages plays into the rest of this commentary.

Even glancing at the first pages, I couldn’t help but arrive at a couple of conclusions:

- **Bad news:** For fair use and balanced copyright. Google’s defense of the suit could have clarified and broadened fair use provisions. I believe Google had a good case. By settling and, in effect, licensing its uses, Google makes it more difficult for another group to claim fair use in digital indexing—and strengthens the hand of those who want everything to be licensed. The settlement may not be adjudication, but it still has precedential qualities.
- **Good news:** For Google. Not only does it get rid of the annoying lawsuit, but it also appears to gain an enormous advantage over others who might wish to enter this space.
- **Good news:** For Big Publishing—that is, the AAP. The big traditional publishing houses may

be in trouble in other areas, as broader and more innovative publishing reduces their market share, but getting Google to settle can't help but be good for AAP.

- **Mixed news:** For authors (except perhaps members of the Authors Guild), readers, libraries and pretty much everybody else.

How does that mixed news play out—and what does the settlement say? I won't comment on the entire settlement or all of the commentary, but I'll tackle some secondary sources and point you to others.

If you just need the facts, read the first section below and go to the ALA OITP site on the settlement ([wo.ala.org/gbs/](http://wo.ala.org/gbs/)) for more.

## The Short Version

Here's how the ALA Office for Information Technology Policy summarizes the settlement in its "2-page super simple summary" ([wo.ala.org/gbs/wp-content/uploads/2009/01/gbs-2-pager-final.pdf](http://wo.ala.org/gbs/wp-content/uploads/2009/01/gbs-2-pager-final.pdf)), edited and reformatted slightly, partly to substitute "OP" for the cumbersome "in-copyright, not commercially available" description of out-of-print books protected by copyright.

- The settlement would end the copyright infringement lawsuit that AAP/Authors Guild brought against Google in 2005.
- Google will continue scanning in-copyright books from library collections into its search database; publishers and authors agree to not sue; Google will continue to enable users to search the full content of the scanned books.
- Google will display up to 20% of an OP book's text (currently only three snippets per book are viewable); previews are different for fiction and non-fiction books; no text display is allowed for some types of books (e.g., anthologies of drama); some books display only "fixed preview" (e.g., dictionaries); users *cannot* print out or copy-and-paste any of the preview displays.
- Google will earn money through advertising and by selling access to the full text of OP books; Google keeps 37% of the generated revenue and distributes 63% to rightsholders (publishers & authors) through a mechanism called the Books Rights Registry (BRR); Google pays \$45 million up front to the BRR for previous scanning.
- Individual users can purchase online access to the full text of OP books through an account with Google; rightsholders or Google will set the price of a book; users have perpetual online access to view the entirety of a purchased book.
- A user can copy-and-paste up to 4 pages of a purchased book with a single command, and can print up to 20 pages with a single command; with multiple commands, a user may copy-and-paste and print the entire book; on printed pages, Google will place a watermark with encrypted identifying information that identifies the authorized user.
- Google will provide free Public Access Service (PAS) to each public library and not-for-profit higher education institution that requests it; a user sitting at a PAS terminal will be able to view the full text of all books in the Institutional Subscription Database (ISD); the ISD generally corresponds to OP books.
- A user can print pages of material viewed on the PAS terminal for a "reasonable" per-page fee set by the BRR; the user will not be able to copy-and-paste text accessed through the PAS.
- Google will sell access to the ISD to universities [and other institutions]; users (faculty, students, staff, researchers, librarians, and others) authorized by the subscribing institution will be able to view the full text of all the books in the ISD; access will continue only for the duration of the subscription; the same copy-and-paste and print options that were available to users purchasing individual access are available to authorized users of the ISD; authorized users can make books in the ISD available to other authorized users through hyperlinks, etc. for course use such as ereserves.
- Google and the BRR will set the price of the ISD; pricing will be based on the number of fulltime equivalent (FTE) users; Google may subsidize the purchase of the ISD for some types of participating libraries; Google may charge a lower price for a discipline-based subset of the ISD.
- The settlement creates four categories of partner libraries that contribute books to the Google book scanning project with different rights and responsibilities: fully participating libraries, cooperating libraries, public domain libraries, other libraries.
  - A **fully participating library** signs an agreement with the BRR, releasing the library from liability for copyright infringement provided the library follows particular rules. The library provides Google with in-copyright books for scanning, and will receive in return a digital copy of each book it

provides; the library may use its library digital copy (LDC) to create a print replacement copy of a book in its collection that is damaged, destroyed, deteriorating, lost or stolen, or to overcome obsolete formats; the library may provide special access to the LDC to a user with a print disability; the library may permit faculty and research staff to use five pages of any book in the LDC that is not commercially available for personal scholarly and classroom use (if the library keeps track of such uses and reports them to the BRR). Fully participating libraries must meet the requirements of the Security Standard (including issues of identification and authentication, access control, network security, risk assessment, and other provisions). Prohibited uses of the LDC include sale of access, interlibrary loan, e-reserves, course management systems, or any infringing uses.

- A **cooperating library** provides in-copyright books to Google for scanning, but does not retain digital copies of the in-copyright books provided by Google. Cooperating libraries do not have to comply with the Security Standard. Cooperating libraries receive a release from any copyright infringement liability if they destroy any past in-copyright digital copies provided by Google.
- A **public domain library** provides only public domain books to Google, and receives a release from any copyright infringement liability if it destroys any past in-copyright digital copies provided by Google; does not have to comply with Security Standard
- **Other libraries** are libraries that have agreed to provide Google books to scan, but have chosen not to participate in the settlement.
- The settlement is non-exclusive: it does not restrict participating libraries from engaging in other digitization projects outside of the Google settlement
- Some participating libraries may be allowed to permit users to conduct non-consumptive research (e.g., linguistic analysis over large collections of textual works) if the libraries agree to specific access and security provisions.

- Google agrees that within five years of the settlement, it will provide free search, the Public Access Service, and institutional subscriptions for 85% of the OP books it has scanned; Google must use “commercially reasonable efforts” to accommodate users with print disabilities.
- The settlement does not apply to books published after January 5, 2009; qualifying rightsholders have until May 5, 2009 to opt out of the settlement class; after May 5, 2009, the U.S. District Court in New York will conduct a hearing to consider the fairness of the settlement.
- Rightsholders who do not opt out of the settlement have until April 5, 2011 to request the complete removal of a specific book from the database.

That really is a super-simple version and it's more than 1,000 words (although a bit shorter than OITP's original). We're dealing with huge, complex documents and agreements here. If you've followed GBS and GLP in the past, you should have already wondered about some of those clauses—e.g., libraries being required to *return* digitized copies they already received, if they don't agree to Google's (revised) terms.

#### *A few more points from Band's guide*

How complex is it? Jonathan Band's *Guide for the Perplexed: Libraries & the Google Library Project Settlement* ([wo.ala.org/gbs/wp-content/uploads/2008/12/a-guide-for-the-perplexed.pdf](http://wo.ala.org/gbs/wp-content/uploads/2008/12/a-guide-for-the-perplexed.pdf)) is 23 pages long (admittedly double-spaced). The settlement documents run to more than 200 pages (they're also available from the ALA OITP site). OITP's list of links to blog posts and articles runs to three pages and more than eighty links—and it's far from complete (but OITP's clearly making a real effort here, as the list includes very recent posts and articles).

As in previous stages of Google Library Project, Band's analysis is readable and thoughtful—well worth reading on its own merits. It's primarily a summary, not an exploration of policy issues. I won't attempt to summarize it further, but will note a few significant items that don't appear explicitly in the 2-page supersummary:

- The preview rules for OP nonfiction books (for general web use) allow multiple five-page sections—but not *adjacent* five-page sections.
- While the general web preview rules for OP *fiction* books allow either 5% or 15 pages at a time (whichever's less), the last 15 pages or 5% (whichever's greater) are always blocked—which seems sensible for fiction: You can't peek at the ending. (In both cases, you can't look at

more than 20% of a book in total—although there might be ways around that, with multiple computers at multiple IP addresses.)

- Except through special publisher arrangements, you'll see *no* context for in-print books (even where "in-print" means "still available through PoD," which is increasingly likely for books): The snippets are gone. Searches may work but you'll just know that the text appears *somewhere* in the book, with no context. That's a significant step backward.
- "Purchase" price for OP books will be "designed to find the optimal price for each book to maximize the revenue for the rightsholder"—and involves "bins" ranging from \$1.99 to \$29.99. But "purchase" is a tricky word here. You're *not* buying the book—you can't download it as a book to a mobile device, sell or lend it to someone else or any of those things. You're buying "perpetual" online access, where "perpetual" may mean "as long as Google is around."
- There's language about providing fully-participating libraries with digital copies of *other* books—books in a library's collection that Google scanned from another fully-participating libraries—that require Google to first scan either 30% of the library's collection or 300,000 books, whichever is larger.
- Fully-participating libraries *can't* create print replacement copies if they can find unused copies of the books for sale "at a fair price." They *can* provide their own finding tools for the digital copies—but can only provide snippets for results. (That may be oversimplified: Can a library use its own finding tools to provide access to the views that would otherwise be available through Google or through licensing?)
- There may be two research centers (outside of Google) holding copies of the entire Google Library Project, at sites selected by fully participating and cooperating libraries and with use limited to "non-consumptive research"—basically statistical and linguistic analysis and data mining that doesn't involve significant actual reading. On the other hand, copyright owners may ask that any book be removed from these databases.
- I added a parenthetical note to the 2-page Super Summary note about selling subscriptions to universities: In fact, Google can (and apparently will) offer such subscriptions to public libraries, K12 and possibly other institutions as well.

So where does that leave us, before considering some of the copious commentary?

- There's no discussion of orphan works—the presumption is that all 70% of scanned books that are OP *have* known rightsholders who can be contacted.
- For far too many books, publishers can limit access beyond its current limits by retaining print-on-demand "in print" status. That's also a problem for authors, since most reversion clauses don't take effect until a book is actually out of print—but that's a more general problem.
- For in-print books where publishers haven't made separate agreements, this agreement ends *any* context for search results. That's a significant loss and an odd one, as you'd think publishers would recognize the usefulness of those snippets for adding new sales. (I know, I know: I'm suggesting that publishers look beyond their noses. Not likely, with the exception of Baen Books and a few others.)
- This agreement puts Google in the licensed-database business for libraries in a potentially big way—and also in a different business, licensing access to individual books by individual users under the misleading "purchase" name.
- Normal users will be able to see a little more of most OP books and a little less of most in-print books—but with less ability to print or cut-and-paste than we've had so far.
- I find it hard to get excited about the one-locked-terminal-per-library "full" access to OP books; it seems like an almost-meaningless sop to cries for increased access. (I could be entirely wrong on that.)
- It strikes me that neither Google nor users will be all that happy with the single-book "purchase" provision, given the quality of some Google scans. Will a user who's forked over \$2 to \$10 be happy when some pages are smudged or unreadable?

### *Putting on several hats*

I'm looping back after having assembled the first draft of this piece, which means I've now read at least 50,000 words of documents and commentary (and assembled a sizable portion of that, including my own interspersed comments). Here's a *personal* take, wearing each of several hats in sequence:

- **As one who cares about fair use:** The settlement saddens me, because I believe Google had an excellent shot at winning the court case and

clarifying fair use. I can certainly understand why Google chose not to do so, but I'm still disappointed—and I believe it will make things more difficult for the next group that wants to make effective and robust use of fair use rights.

- **As a library person:** Mixed feelings. Continued availability and growth of Google Book Search: A good thing, and good for libraries. Expanded preview visibility for OP books: A very good thing—and also good for libraries. The one free terminal: Largely irrelevant except for very small libraries, in my opinion—it really does have the feel of a “the first one is free” offer. Individual “book sales” that are really licenses for “perpetual” online access to page images not always terribly well scanned: I have trouble getting excited about this, either as a service to the public or as competition for libraries. Subscriptions to the full-book database: Also mixed feelings, depending on price and other issues. Having five million more “books” is probably a good thing; having another chunk taken out of too-small budgets, not so much.
- **As one who's cautious about Google's power:** I'm not thrilled, particularly because—while this settlement *in no way* “privatizes” library collections (I'm getting to really hate that formulation)—the realities do tend to give Google a de facto monopoly over very large scanned OP provision.
- **As a book reader:** I see myself using Google Book Search and the free view capabilities. I don't see myself “buying” “books” this way. But who knows? I certainly see this as continuing to be a good way to find books I want to borrow, in physical form, from a library.
- **As an author:** Good, good, unclear. **Good:** My OP works will be more discoverable and more viewable—and yes, I've gone in, claimed the (9 to 18, depending how you count) OP books where rights have reverted to me and asserted that my Lulu books *are* commercially available (but left preview provisions in place). **Good:** I'll be happy to take \$60 times some small number, when things finally play out. (One OP book should now have rights controlled by RLG; it's up to OCLC to make that claim. For one ALA Editions book, I'm not sure whether it's actually OP, so don't know whether rights have reverted.) **Unclear:** Whether I can make effective use of the Registry. I've seen a claim

that it requires a \$200 payment to become part of that, in which case I wouldn't. If that turns out not to be true and the actual costs are zero or nearly so, I *will*—and, in doing so, will set the “purchase” price of my OP books at \$0.

Confusing enough? Read on. Let's look at what some others have had to say.

## Very Early Commentaries

The proposed settlement was announced on October 28, 2008. Commentaries began immediately. I'm excerpting a few of many—some wholly in favor of what they saw, a few mostly opposed, many with conflicting feelings.

### *Disruptive library technology jester*

That's Peter Murray, blogging at [dlj.org/](http://dlj.org/). He quotes portions of a plaintiff's motion that show how the big publishers and one group of authors view this:

[T]he Settlement:

- Creates an innovative marketing program for authors and publishers of in-print books that catapults the publishing industry into the digital age, a result that greatly benefits individual authors and publishing houses, which simply could not launch such a program on their own;
- Addresses what has been a persistent problem, particularly for individual authors—how to breathe new life into older, out-of-print books that are generally inaccessible to the public and have stopped generating revenue;
- Is designed to maximize Settlement Class member rights by allowing any of them, at any time, to commercially exploit their works in other ways outside of the Google Library Project; and
- Benefits the Settlement Class, as well as the general public, through the ability to access books on Google's website and, as a result of provisions addressing the extent to which libraries may also use digitized copies of these works, enjoy a new and unprecedented ability to use books and conduct research....

[T]he result is a settlement that, although complex in its structure, is elegantly simple in its result. It provides extraordinary and previously unattainable benefits to the authors, the book industry, and even the public.

Marketing, maximized rights for authors, more revenue from OP books—oh, yeah, and there's a “general public” benefit too, sort of. I think AAP and AG get it right here: This settlement *primarily* benefits (some) authors and (some) publishers.

The next day, Murray reviewed the *Notice of Settlement* itself (only 38 pages, not the full 200+). Leav-

ing out material already covered, a few items Murray points out:

- Photographers and illustrators aren't covered—and that may include book jacket designers.
- Books are only covered if they're actually registered with the Copyright Office, which hasn't been a requirement for copyright since 1976. (As explained elsewhere, this is probably the only way a court can handle it.)
- Periodicals aren't covered—and there are certainly periodicals in GBS.
- The Book Rights Registry itself will take 10-20% of revenues as an administrative fee.
- It's going to take some time: "The parties warn that it is going to take a while to make this happen. It will take considerable time to implement the commercial uses authorized under the Settlement, implement the elections made by Rightsholders for their Books and Inserts, and make Cash Payments.... Please be patient, and visit the Settlement Website at [www.googlebooksettlement.com](http://www.googlebooksettlement.com) regularly for updates."

In other posts, Murray looks at the Public Access Service, how the settlement affects library consortia, and the preliminary court approval (noting that this settlement seems to be proceeding with surprising haste, given the long prelude). Actually, the settlement refers to "Institutional Consortium" but with a strikingly library-specific definition:

"Institutional Consortium" means a group of libraries, companies, institutions or other entities located within the United States that is a member of the International Coalition of Library Consortia with the exception of Online Computer Library Center (OCLC)-affiliated networks.

One could reasonably ask what constitutes "OCLC-affiliated" but let's not. As Murray analyzes the special provisions related to consortia, he concludes that one provision currently boils down to a single consortium: CIC. Other provisions aren't quite so specific. (I've included notes on *DLIJ* commentaries from November 2008 and beyond. Peter Murray did yeoman service in getting these points out early and clearly.)

### *The distant librarian*

That's Paul Pival, offering a Canadian perspective on October 28 ([distlib.blogspot.com/distlib](http://distlib.blogspot.com/distlib)). As he notes, the current settlement doesn't directly affect Canadians or anyone else outside the U.S.—but that could change. (Probably not for the better: Can Google justify showing snippets of in-print books to Canadians when it can't show them to U.S. viewers?)

Pival wonders about Google's comment that its current service offers results "like a card catalog": "Is a **card** catalogue really the analogy most of your users are going to understand these days?" He comments on the apparent benefits for OP books, which Google calls "volumes they might have thought were gone forever from the marketplace": "Yeah, like the publishers couldn't reprint volumes they thought were gone forever? And let's not mention the tremendous boon to people who want to read the books!" That first sentence is tricky—with good reversion clauses, the publishers no longer own the rights to those books. The key here, though, is that Google's providing potential revenue *with no effort on the part of authors or publishers*, as opposed to even the minor effort of setting up a PoD offering of an OP book.

Pival regards this as "really great stuff." At that point, he didn't appear to see much downside.

### *Confessions of a science librarian*

John Dupuis posted "The Google Books Search deal: A real game-changer" on October 29, 2008 at [jdupuis.blogspot.com/](http://jdupuis.blogspot.com/). He focuses on the licensing and purchase provisions—and notes that he sort of predicted this possibility a few years ago. A key paragraph:

I can't wait to see details on this, especially if there will be some sort of DRM, how printing will work, whether or not you'll be able to download to readers such as the Kindle. Of course, it will be really interesting to see what a site license for a large university will cost. Will it be the equivalent of our entire monograph budget? The implications and the choices that would imply are staggering. Talk about a rock and a hard place. This has the potential to completely transform the ebook business and the way libraries buy books. The traditional players in the ebook business will have to really focus on seriously adding value to their offerings, the way A&I services have to add more value in the face of Google Scholar. Libraries will be faced with a lot of choices, especially in the face of fears of putting all our eggs in one basket.

From what I can tell so far, it appears there *will* be some sort of DRM and you *won't* be able to download to ebook readers. Beyond that, we just don't know.

### *Open access news and linked items*

Peter Suber provided his own early thoughts (at [www.earlham.edu/~peters/fos/](http://www.earlham.edu/~peters/fos/)) and linked to lots of other commentaries. I think Suber's first impressions are so much on target they're worth quoting extensively:

*What looks good here?* Google will continue to scan copyrighted, OP books (as well as public domain books) and make them full-text searchable. Those searches will continue to be free of charge and may

now display much more than short snippets... Publishers are dropping their objection to future scans, which will encourage more libraries to participate in the program and enlarge Google's book index. Publishers of non-OA books have found a way to enter the 21st century without shunning the internet or losing money.

*What looks bad here?* Other book scanners may have to pay to play as well, even if Google's original fair-use claim was valid. The settlement may reduce scanning of copyrighted books by everyone except Google.

Some of Google's \$125 million will set up the Book Rights Registry and some will be "compensation" to publishers whose books have already been scanned... I can't tell whether Google will "compensate" publishers for future scans or merely share revenue with them. That may look like a fine point. But if Google will compensate publishers for future scans, then it has relinquished its fair-use claim: that the scanning was lawful without permission or payment provided the company displayed only short snippets. But if Google is merely sharing revenue, then it hasn't necessarily relinquished that claim. Giving up a valid fair-use claim would be a serious loss and could tie the hands of search engines forever. Moreover, the claim seemed valid to a gaggle of copyright specialists...

See our many past posts on this lawsuit and my article from October 2005, "Does Google Library violate copyright?" In that article I called the publisher lawsuit a shakedown, and so far I see no reason to change my mind.

Suber heard from Google's Derek Slater noting that Google would be compensating rightsholders for past scans with \$60 per book—and asserting that they would *not* be paying compensation for future scans, "though we will have a revenue share..." I read that as a de facto relinquishing of the fair-use claim. As for Suber, he added more thoughts in a November 6, 2008 post after Seth Finkelstein said in a *Guardian* column that the settlement would have OA advocates up in arms. Excerpts:

- I'm deeply disappointed that Google didn't litigate the fair-use claim to the end... (1) Google had a strong case, (2) almost nobody else could bear the enormous legal costs of fighting the AAP and AG, (3) the proposed settlement weakens the claim for any future litigant, if only by creating a new commercial opportunity for publishers to balance against fair use, and (4) leaving the fair-use claim unresolved is harmful to digitization projects and search engines. So yes, I'm up in arms about that aspect of it.
- On the other hand, I'm not at all sure that litigating the claim to the end would have been a victory for Google and fair use. . As I wrote in a 2005 article,

"On the merits, it's an important question to [resolve]. But I admit that I'm not very comfortable having any important copyright question [resolved] in today's legal climate of piracy hysteria and maximalist protection..."

- Google and the publishers disagreed passionately about the fair use claim, each side thought it was right on the merits, and each wanted to see the question resolved in its favor. The settlement must have been delayed by the fact that neither side could readily give up the legal claim it thought was so essential to its business. But both sides understood that fair use is vague and contestable, and neither wanted to take the risk of seeing the claim resolved the other way. Choosing to settle instead is a hard judgment but, in the end, I'm not sure it was wrong. **The settlement will harm fair use, but refusing to settle might have been more harmful.** This consideration vents much, but not all, of my steam. [Emphasis added.]
- Harvard is right that the settlement puts needless restrictions on the digitized editions of the copyrighted, OP books at the heart of the case. But we have to remember that we wouldn't have OA to these texts even if Google had not settled and had prevailed in court on every point. In that sense, the major issue isn't OA at all, but what sort of restricted access different kinds of users would have to books...under copyright but out of print. That's an important question, for research and commerce, but it doesn't implicate OA. You might wish that OA had been an option, but the OA movement deliberately focuses on works which pay no royalties, like journal articles and public-domain books, or works with consenting rights holders. From this point of view, OA itself was not at stake in the lawsuit or the settlement.
- By contrast, OA really is at stake for Google's digitization of public-domain books. But neither the lawsuit nor the settlement affects that part of Google's project, and Google plans no changes to it. On that front, btw, I've always argued that Google's form of access is a step forward which stops short of OA, and that the OCA model is superior precisely because it is OA.
- I do hope that Harvard can persuade the parties or a court to tweak the restrictions on the copyrighted books at the heart of the case. But the fact that the settlement itself could be improved doesn't change the fact that, even unmodified, it would be an improvement over the kinds of access we have to copyrighted books today: 20% previews rather than short snippets, free full-text access from selected terminals in libraries, free text-mining of full-texts for some institutional users, free full-text searching of a larger rather than smaller number of books, and even the availability of priced access to full-text digital editions. If OA were an option, I wouldn't be nearly as

happy with these half measures. But for these copyrighted books, it was never an option.

Suber's first roundup (on October 30) includes an impressive array of opinions. Some excerpts (I may not be choosing the same excerpts that Suber did):

- From Andrew Albanese's October 28, 2008 *Library Journal Newswire* article, "Google settles landmark lawsuit over book scanning":

On a conference call this morning, the parties said that there remained a strong difference of opinion over the copyright principles at the core of the case. "We had a major disagreement with Google, and we still do," said Paul Aiken, executive director of the Authors Guild. "We also don't see eye-to-eye on with publishers on book contract law," he added, before calling the settlement the "the biggest book deal" in U.S. publishing history. Aiken said two "guideposts" helped lead his organization through a thicket of issues in the suit. "Authors like their books to be read," he noted, "and they like a nice royalty check."

- From Kirk Biblione at *Medialoper* ([medialoper.com/](http://medialoper.com/)):

At first glance, this appears to be the rare settlement agreement that seemingly benefits all parties. In fact, the only entities that don't seem to have fared so well are parties who weren't involved in the suits.

### The Winners

- **Google:** It's hard to overstate how important this agreement is for Google. Google has essentially acquired the digital rights to the long tail. At least the portion of the long tail that's locked up in out of print books. ... Google has mastered the art of turning arcane search phrases into money. In the future they'll have a lot more content to monetize. Content that no other search engine will have access to. That's a huge competitive advantage.
- **The Rightsholders:** Authors and publishers will benefit immediately as they allocate the funds from the initial settlement, and over time as they collect revenue generated from out of print works. In the vast majority of cases, these out of print works would have never generated any additional income... Google has basically created an entirely new revenue stream that publishers can use to profit on books that would otherwise not have generated a cent.
- **Libraries:** The libraries that participate in the digitization program will get to keep control over their archives. Equally important, libraries will have digital access to the archives of other libraries. The academic community as a whole will benefit in ways that we can't yet imagine.
- **The Public:** The public gets easy access to millions of rare and out of print works.

### The Losers

- **Amazon:** Amazon's 190,000 Kindle titles look puny compared to the millions of books Google now has access to. Granted many of those Kindle titles make up the big head of consumer demand, as opposed to the long tail. Still, Google now has the ability to monetize millions of books Amazon can't, if for no other reason because they're out of print. What's more, under the new agreement Google has the right to sell printed copies of those books via print on demand. And I have a sneaking suspicion that Google still has a few more surprises in store for us.
- **Microsoft:** Not long ago Microsoft had its own book search program. The company unceremoniously killed that program on the eve of BEA earlier this year...
- **Fair Use Advocates:** There are many (myself included) who believed Google had a strong fair use argument to support their scanning efforts. It was hoped that a Google court victory would reaffirm those rights. By settling out of court Google avoided the issue entirely...

I'd argue that this analysis oversimplifies effects on libraries and may exaggerate public benefits, and omits the extent to which OCA and similar projects may be losers. I would note that, for many OP books, the "rightsholder" should *not* be the publisher—but will most authors know to sign up for the registry?

- Paul Courant at *Au courant* ([paulcourant.net/](http://paulcourant.net/)):

First, and foremost, the settlement continues to allow the libraries to retain control of digital copies of works that Google has scanned in connection with the digitization projects. We continue to be responsible for our own collections. Moreover, we will be able to make research uses of our own collections. The huge investments that universities have made in their libraries over a century and more will continue to benefit those universities and the academy more broadly.

Second, the settlement provides a mechanism that will make these collections widely available. Many, including me, would have been delighted if the outcome of the lawsuit had been a ringing affirmation of the fair use rights that Google had asserted as a defense... But even a win for Google would have left the libraries unable to have full use of their digitized collections of in-copyright materials on behalf of their own campuses or the broader public... Making the digitized collections broadly usable would have required negotiations with rightsholders, in some cases book by book, and publisher by publisher...

The settlement cuts through this morass. As the product develops, academic libraries will be able to license not only their own digitized works but everyone else's.



Michigan's faculty and students will be able to read Stanford and California's digitized books, as well as Michigan's own. I never doubted that we were going to have to pay rightsholders in order to have reading access to digitized copies of works that are in-copyright. Under the settlement, academic libraries will pay, but will do so without having to bear large and repeated transaction costs. (Of course, saving on transaction costs won't be of much value if the basic price is too high, but I expect that the prices will be reasonable, both because there is helpful language in the settlement and because of my reading of the relevant markets.)

Courant's post drew several dozen comments, many of them less sanguine about the situation. It's a fascinating discussion.

➤ James Grimmelmann at *The laboratorium* (laboratorium.net), excerpts:

**This is a Google-only deal.** The result of the settlement will be to give Google a license to keep on doing what it's doing, while allowing the authors to use their now-sharpened knives to sue anyone else who tries to do the same. At that point, of course, Google would be delighted for the authors to succeed, since it keeps the competition at bay. The settlement may also be bad for other search engines in another respect: the authors will claim that it undermines any claim of fair use in indexing books and making them searchable...

**You can't strike a deal like this without court approval.** That matters, because even if this settlement is approved, there is still no functioning "market" for these uses of copyrighted works. The issue is that this is a class-action settlement requiring judicial approval to bind all authors. It's practically impossible for anyone else to take advantage of Google's terms without filing suit to obtain a similar class-binding order...

➤ Lawrence Lessig at *Lessig 2.0* (lessig.org/blog), brief excerpts from a fairly long post:

IMHO, this is a good deal that could be the basis for something really fantastic... [T]he settlement does not presume to answer the question about what "fair use" would have allowed. The AAP/AG are clear that they still don't agree with Google's views about "fair use." But this agreement gives the public (and authors) more than what "fair use" would have permitted. That leaves "fair use" as it is, and gives the spread of knowledge more that it would have had.... The biggest loser in this whole battle is the Orphan Works legislation. If anyone needed evidence to demonstrate that it is way *too early* for Congress to be passing massive new bureaucratic overlays to copyright to deal with the important problem of "orphan works," this is the evidence. Let's let this private alternative develop, while Congress puts away its billion-factor balancing tests for regulating access to "orphan works"...

Lessig has moved on from a copyright focus, but it's still sad to see him scare-quoting fair use and orphan works. The comments are worth reading.

➤ Sherwin Siy at *Public Knowledge* (www.publicknowledge.org), brief excerpts:

There's a lot to be debated in this settlement...but let's first note what it *doesn't* do: make a determination as to what is or isn't fair use. Depending on how you saw the merits of the case, and how confident you were in the court reaching the right decision, that can be good or bad. On the one hand, we don't have a federal court saying that scanning books is a per se fair use; on the other hand, we don't have a court saying that scanning is per se infringement, either. This does mean that the financial and legal might of Google is no longer going to be aligned with libraries and archives that may wish to provide digital services that are technologically similar to Google's efforts. This will mean that further fair use fights for digital libraries start closer to square one than they would have otherwise... But while the legal landscape isn't altered too much by the settlement, the practical landscape could be. Rightsholders and other potential plaintiffs might view this settlement as the model for all future relationships with digitization efforts—if Google pays for digitizing, why shouldn't everyone else? Such a landscape might make a plaintiff more likely to sue, although the results in court, ideally, shouldn't differ, with or without this settlement in place.

➤ Mike Masnick at *Techdirt* (techdirt.com), brief excerpts:

Pretty much any way you look at it, Google caved here--and this is unfortunate for a variety of reasons.... Two years ago, there was a story in the *NY Times* about how Google's legal department saw all of these lawsuits against the company as a way to stand up on principle and make better law. Specifically, the company positioned itself as being willing to fight certain lawsuits on principle in order to get precedent setting rulings on the books in support of openness, fair use, safe harbors and many other important issues.... [I]t's quite upsetting to see Google cave on this. The settlement does not establish any sort of precedent on the legality of creating...an index of books, and, if anything pushes things in the other direction, saying that authors and publishers now have the right to determine what innovations there can be when it comes to archiving and indexing works of content. Unfortunately, this was really inevitable. As was the case with Google caving on YouTube and the Associated Press, it becomes a situation where Google realizes it can throw a little cash at the problem to make it go away--while also creating a large barrier to entry for any more innovative startup. From a short-term business perspective this might make sense, but from a long-term business perspective (and wider

cultural perspective) it's terrible. It will only encourage more lawsuits against Google for trying to innovate, as more and more people hope that Google will settle and throw some cash their way. Furthermore, it greatly diminishes the incentives for making books more useful, and that's damaging to our cultural heritage. While it was always silly to believe that Google ever really operated on a higher principled stance, rather than a short-term business focus, this settlement is tremendously disappointing.

- Suber gave Harvard its own entry, "Harvard doesn't like the Google settlement." While we'll see more of Robert Darnton's comments later, here are a few excerpts from the *Harvard Crimson* article quoted by Suber:

Harvard University Library will not take part in Google's book scanning project for in-copyright works after finding the terms of its landmark \$125 million settlement regarding copyrighted materials unsatisfactory, University officials said yesterday... University officials said that Harvard would continue its policy of only allowing Google to scan books whose copyrights have expired...

University spokesman John D. Longbrake said that HUL's participation in the scanning of copyright materials was contingent on the outcome of the settlement between Google and the publishers. Harvard might still take part in the project, Longbrake said, if the settlement between Google and publishers contains more "reasonable terms" for the University.

In a letter released to library staff, University Library Director Robert C. Darnton '60 said that uncertainties in the settlement made it impossible for HUL to participate. "As we understand it, the settlement contains too many potential limitations on access to and use of the books by members of the higher education community and by patrons of public libraries," Darnton wrote. "The settlement provides no assurance that the prices charged for access will be reasonable," Darnton added, "especially since the subscription services will have no real competitors [and] the scope of access to the digitized books is in various ways both limited and uncertain." He also said that the quality of the books may be a cause for concern, as "in many cases will be missing photographs, illustrations and other pictorial works, which will reduce their utility for research and education."...

The comments on the *Crimson* site are fascinating ([www.thecrimson.com](http://www.thecrimson.com)), including one person who argues that anyone publishing *any* nonfiction book should be *legally required* to deliver an analytical index, "meeting the highest standards," in electronic format to a site such as LC, prior to publication. Not relevant to this discussion, but wow! So much for my

self-published books: **Illegal**, every one of them, if this Harvard person had his way.

### *Academic librarian*

Wayne Bivens-Tatum posted "What's so bad about Google?" on October 31, 2008 ([blogs.princeton.edu/librarian](http://blogs.princeton.edu/librarian/)). Excerpts:

The Google Books settlement seems like very good news to me. I'm assuming the cost of subscribing to the full Google Books service won't be prohibitive for most libraries, and that means that a lot more people will get a lot more access to even copyrighted books. The only way it could be better for me is if the Google-scanned books were in Mobipocket format as well as PDF so I could read them on my phone. Small quibble, though.

Nevertheless, the news stories have managed to find some critics of the plan. This criticism I found the most compelling:

"I will tell you, frankly, that I kind of wish this case had gone to litigation. I think Google had a great fair use defense," agreed Corynne McSherry, staff attorney for the Electronic Frontier Foundation, which advocates internet free-speech rights, in the October 29 San Francisco *Chronicle*. "A ruling from the court would have been good for everyone. It potentially could have fostered other offerings, based on that legal certainty" that would have stemmed from a Google win.

I found it compelling because I didn't agree with the publishers that the scanning was an abuse of copyright... However, given the current rage for draconian copyright policy, the courts would probably have tried to shut the Books project down completely. No use is fair use in the eyes of the publishers.

Other criticisms are leaving me cold at the moment. This article (found via LISNews) has one such criticism from Brewster Kahle:

"When Google started out, they pointed people to other people's content," Kahle said. "Now they're breaking the model of the Web. They're like the bad old days of AOL, trying to build a walled garden of content that you have to pay to see."

Breaking the model of the Web? One might wonder if the Web is really old enough to have a model that it would be indecorous to break... What precisely is wrong with a "walled garden of content that you have to pay to see," when the alternative was a walled garden of content that you'd have to pay to see? Prior to this deal one would still have to find a library or purchase a book to read the whole thing. Now you'll be able to do that online, and probably for less money. True, Kahle and the Open Content Alliance want all these copyrighted books freely available on the Web

for everyone. But, as the article noted, “they haven’t figured out how to make it work.”

The New York *Times* noted the following criticism:

“On the one hand, one admires all of Google’s inventions,” said Rick Prelinger, board president of the Internet Archive, a nonprofit organization that has scanned and made available online one million public domain books. “But when you start to see a single point of access developing for world culture, by default, it is disturbing.”

Perhaps the problem is me, but I just don’t understand the criticism. A single point of access for world culture? I suppose Google would like that, but is that what’s happening with the Google Books project?. Some critics are acting like the only place to find any of these books is Google itself. Google isn’t publishing the books. They’re still accessible through libraries and bookstores. And for the ones out of copyright, there are already plenty of other places to get many of these books, only not on the scale of Google.

That might be what some people don’t like, the scale of Google that other groups haven’t been able to achieve. The Open Content Alliance wants to make everything freely available, but they haven’t figured out how to do it. The Internet Archive has scanned a million public domain books, but Google’s done more and still made them freely available. The criticisms just seem like sour grapes to me...

I would take issue with portions of this. First, as far as I can see, Google-scanned OP books won’t be available for download *period*—in PDF or any other form. Second, OCA has *never* said all copyright books *should* be freely available on the web (as far as I know), although they’d like to improve access to orphan works. To date, OCA has entirely concerned itself with public domain works. Third, much of the “single point” issue is that this agreement seems likely to make it more difficult for anybody else to do similar scanning.

### ***Electronic Frontier Foundation***

Fred von Lohmann posted “Google Book Search settlement: A reader’s guide” at *Deep links* ([www.eff.org/deeplinks/](http://www.eff.org/deeplinks/)) on October 31, 2008. Excerpts:

So far, two things are plain.

First, this agreement is likely to change forever the way that we find and browse for books, particularly out-of-print books... Second, this outcome is plainly second-best from the point of view of those who believe Google would have won the fair use question at the heart of the case...

But the settlement has one distinct advantage over a litigation victory: it’s much, much faster. A complete victory for Google in this case was probably years

away. More importantly, a victory would only have given the green light for scanning in order to index and provide snippets in search results; it would not have provided clear answers for all the other activities addressed in the settlement, such as providing display access for out-of-print books, allowing nondisplay research on the corpus, and providing access for libraries...

It seems likely that the “nondisplay uses” of Google’s scanned corpus of text will end up being far more important than anything else in the agreement. Imagine the kinds of things that data mining all the world’s books might let Google’s engineers build: automated translation, optical character recognition, voice recognition algorithms...

This agreement promises unprecedented access to copyrighted books. But by settling for this amount of access, has Google made it effectively impossible to get more and better access? The agreement allows you to “purchase” digital access for out-of-print books, but does not include the right to download the book (unlike public domain books). So you can read the book, but only on Google’s terms. Libraries get more access, but for an undisclosed price...

...If Google becomes the default place to search, browse, and buy books, it will be able to keep unprecedented track of what you read, how you read it, and collate that with all the other information it has about you. Does the agreement contain ironclad protections for user privacy?

The post raises other concerns that von Lohmann was considering as he analyzed the complete agreement.

## **Later Commentary**

Most of the previous items appeared within three days of the proposed settlement. Others weren’t far behind. A few more excerpts since then—omitting two major sets of commentaries that began in November 2008 but deserve separate discussion. (Thanks to Peter Suber of *Open access news* for links to some of these items and to ALA OITP for others. Some came from my own blog reading.)

### ***Open Content Alliance***

“Let’s not settle for this settlement” appeared on November 5, 2008 on the OCA website ([www.opencontentalliance.org](http://www.opencontentalliance.org)). Excerpts, not repeating material quoted from other sources (much of which I’ve already quoted):

Rather than accept the Google settlement with publishers and authors as a fait accompli, or as an obligatory blueprint for the future, the appropriate response is to consider its implications for the future and take all steps to build the world we want to live

in. Although the settlement may solve some immediate problems for the parties to the lawsuit, and perhaps some of the contributing libraries who have enabled it, we should not assume that Google Book Search is the only way, or even the best way, to organize and make available our cultural heritage...

Losing access and control of our cultural heritage as part of a digitization wave is not acceptable. At its heart, the settlement agreement grants Google an effective monopoly on an entirely new commercial model for accessing books. It re-conceives reading as a billable event. This reading event is therefore controllable and trackable. It also forces libraries into financing a vending service that requires they perpetually buy back what they have already paid for over many years of careful collection...

...The issues encompassed by the Google-AAP-Authors Guild Settlement extend beyond the interests of the parties to this one lawsuit. We believe that, as a society, we can do better. We encourage you to read the settlement and the various commentaries we've linked to above and then to join us in working toward an open library system for the digital age.

Let's define an appropriate response to achieve our shared goal: Universal Access to All Knowledge.

I earlier defended OCA against the charge that they seek to make *all* works freely available. As bluntly stated at the end of that commentary, maybe I was too quick to defend.

In the selective set of quotations from others, there's a factual error: "Harvard University has chosen not to continue its participation with Google..." In fact, Harvard has chosen not to *expand* its participation to include in-copyright books. The article linked to clearly states that Harvard will continue to allow Google to scan works in the public domain, which was all Harvard had actually signed up for.

Portions of the second paragraph are, at best, questionable, but consistent with Brewster Kahle's general level of assertions. Some forms of reading (e.g., many commercial databases) have *always* been billable and trackable; the settlement neither makes all reading billable nor transforms *any* existing sources of reading into billable units. The settlement **in no way** "forces libraries...to buy back what they have already paid for": The books are on their shelves and *at least* as useful as they were before (I'd assert they are *more* useful because GBS will continue to exist and be freely available). The settlement **in no way** causes anyone to lose access to anything they had access to before; this argument is as absurd as the repeated argument of one critic that libraries are "giving away" their books or somehow "privatizing" them by allowing Google to

scan *and return* them. There's plenty to criticize in the settlement; it shouldn't be necessary to bend the English language out of shape in order to do so.

A month later, "A raw deal for libraries" appeared on the OCA site, starting out "One of the most surprising, even shocking, features of the Google-AAP-Authors Guild Settlement is how hard it is on libraries." This time, there's no question: This is a flat-out attack on the deal. Extensive excerpts:

Given that Google Book Search could not have gotten off the ground without the cooperation of various university libraries, it is particularly disheartening that the proposed settlement treats them with such an iron fist at the same time as it expects them to foot much of the bill through subscriptions. It will be interesting to see how many libraries continue as partners, given Google's bait-and-switch.

Take for example the digital copy that Google gives to a library in exchange for scanning its copy of a book. Previously, all library partners were given digital copies. According to the proposed settlement, however, only "fully participating libraries" will continue to receive copies from Google... All other categories of libraries will no longer receive copies in exchange and, to make matters worse, they will have to destroy the digital copies of in-copyright books they already possess or otherwise expose themselves to the implied threat of a lawsuit from authors and publishers over copyright infringement.

Yet even these "fully participating" libraries are granted only a few permissible uses of their copies...while other uses that are arguably fair use (inter-library loan, use in e-reserves and course management systems) are strictly forbidden... How far we have fallen... Fully participating libraries must now give up [other apparently-contractual] benefits and, if that wasn't sacrifice enough, they must also guarantee the security of their digital copies as laid out in a 17-page "security standard," under the threat of fees up to \$7.5 million for security breaches.

Libraries have made huge investments in the books that Google is digitizing. Not only did they purchase, process, shelve and care for the books, over many years, but they continue to carry significant overhead costs for their continued use (including Google's use!). Much of this investment has been made with taxpayer dollars. And yet libraries receive 0% in this proposed settlement while Google gets 37%. What kind of partnership is this? Taxpayers should be alarmed that their money has gone to provide a service that Google is exploiting on its own terms, in its own interests, with no monetary and little other return to the libraries...

The comments here are lengthy—nearly seven times as long as the post itself, beginning with a response by

Dan Clancy (of Google Book Search) that, all by itself, is more than three times as long as the post, which suggests that the anonymous blogger (Kahle? Rick Prelinger?) struck a nerve. Excerpts from Clancy's response, which begins by calling some of the points in the post "inaccurate" and others "misleading," while emphasizing Google's support of public debate—as long as it's factual:

The settlement agreement opens up new opportunities for reading as it provides explicit authorization that goes above and beyond what would be allowed under fair use. The biggest benefit of the agreement is the fact that the large majority of these books will be accessible in the U.S... While there are many benefits to libraries, the core product offerings are the biggest. For institutions that choose to subscribe, their users will be able to access all of the books in the subscription at no cost to the individual user. For schools that do not have extensive libraries, this should prove very beneficial and even for schools with large libraries this extends the reach of these libraries. However, for schools that choose not to subscribe, their users will still be able to freely preview books and then can choose to purchase the book, access the book through a local library which they can find through the Find It in a Library link or access the book through the access points at public and academic libraries free of charge...

As part of our acknowledgement of [partner libraries'] critical preservation effort, Google offers library partners two options for their own access over and above the other benefits we are offering. For libraries that want access to the entire institutional subscription, Google will pay for a portion or potentially all of the cost of the institutional subscription based upon the number of books scanned from that library. For our partners where we are scanning a large portion of the library, the subsidy is such that these institutions will likely receive a free version of the institutional subscription. This means that for some universities, Google is absorbing the cost of digitizing their entire collection or a large portion of their collection, and in return their students, faculty, staff, visitors and other members of their community will be able to obtain broad electronic access to a large majority of these books as well as access to books scanned from other libraries. For other partners, this subsidy results in a significant reduction in their cost to obtain the subscription to all of the books we digitize from all partners. Alternatively, we also are offering each partner an option called the "limited subscription" that will be free to them. A "limited subscription" provides members of the partner institution access to all of the books that we scanned with them that are included in the larger institutional subscription. Both of these offers extend as long as these books are in copyright and

being offered as part of the program. Once the books are PD, we provide access for free. So, the simple story is: if Google scans a book from a library and is offering it for sale in an institutional subscription product, then that book will be made available to the students, faculty, staff and visitors for that institution at no charge...

[T]he scope of the copies that are returned to the libraries is greater. Under our current agreement, a library partner only receives copies of files from books we scan from that library. In the settlement agreement, for partners that go over certain thresholds of scanning, they will be able to receive copies of books that we scan from other libraries. This ensures that multiple parties will have copies of these files to preserve for posterity...

[Extensive commentary on digital copies and the security-breach issue, best read in the original.]

...It is true that, just like our original agreements, there is no direct revenue sharing in the agreement. Nor is there direct contribution by the libraries to Google's costs. Instead, Google makes a large investment in the digitization of these files to provide many benefits to our library partners, not the least of which is furthering their missions of increasing access to the books including offering free access to the digitized versions to their entire campuses. Perhaps if libraries were for-profit corporations a different deal might have been desirable: one which put money into their pockets and did less public good. We don't believe the OCA members would have liked that deal with the OCA and similarly we believe we have struck the right balance in this deal for our library partners...

...While indexing and search is very beneficial and leads to increased discoverability of books, in the end, most users want to access the books once they discover them, and they want this access to be seamless. Today in Google Book Search, over 70% of the books we have scanned are in "snippet view." While Google is confident we would have prevailed in the lawsuits, we still would have been left with the vast majority of these books inaccessible to users with no clear path to unlocking them. I personally am excited about this agreement because it unlocks an incredible number of books for readers to read and helps to realize the dream of increased access to information.

After an anonymous hardline copyright comment from "A. Writer," Karen Coyle responds to Dan Clancy (at much shorter length) from a library perspective. It's an exceptionally good commentary on the library issues, so I'm including extended excerpts:

...[Y]ou are looking at this from the point of view of a for-profit organization, and that's not a view that includes libraries.

To begin with, you have ignored the question of “fair use.” This is a key aspect of the copyright law that allows the public to make use of copyrighted materials without the permission of the copyright holder. By placing the digital copies behind a subscription service and regulating access and use, fair use (that is, use that is uncontrolled) is not an option. You can be generous, you can be “fair,” but you cannot be the copyright law. This is key because it is the essence of the balance between the commercial interests of the authors and publishers, and the public’s interest in access to the intellectual output of our culture.

As non-profit, educational institutions, libraries enjoy broad fair use rights. They also make use of the “first sale” doctrine. Users can read entire works held by the library without any payment to the rights holder. There have been calls from the rights holders to eliminate this “privilege,” to require libraries to pay to lend books and other materials. Every move in that direction endangers the balance between the rights of the copyright holder and the rights of the public to freely (as in free speech) access information and culture. Without first sale, rights holders retain control over their published works in a way that could easily lead to discrimination and censorship. Open access to materials in libraries is the only defense we have against that. Among other things, this means that items in libraries cannot be withdrawn by the rights holders; no one can go back and revise history. It is this commitment to the public that makes libraries invaluable. With a system in place where everyone pays to view, and where rights holders can potentially recall their works, the rights of the public are no longer being met. Yet, I can easily imagine some cash-strapped communities deciding that they can eliminate their library and just provide access to the Google Book service. That hits me in the gut like a big gulp of 1984.

Next, you ignore section 108 of the copyright law. Section 108 allows libraries to make copies of items in their collections under certain circumstances. Primary ones are: 1) to replace a deteriorating item that is no longer available in the marketplace and 2) to serve disabled users. Both of these are listed as allowed under the agreement with AAP—that libraries can use their copies of the digitized items for these purposes. This bothers me because it puts into a contract something that should be left to copyright law... I see this contract as another erosion of [existing] rights, which means an erosion of the rights of libraries to serve the public as copyright law intends...

[The next paragraph addresses A. Writer’s claim of illegal activity.]

...[L]ibraries in our country are an organic whole, with the actions of a few affecting all. We use the

same standards, we share our resources through inter-library loan, we make broad agreements that benefit the entire community. This is totally unlike the for-profit world where each company looks after itself. We don’t do it that way in libraries, or at least we haven’t done so up until now. I consider the Google partnership with the libraries to be dangerous because it “commercializes” library materials. I know that libraries are impoverished, slow moving, while Google is rich and quick. I would love for libraries to be rich and quick. But in no way do I want them to take on the assumptions or point of view of a for-profit approach to information. Our society would lose so much if that were to happen.

You’ll read more of Coyle’s comments in a later section of this Perspective.

Finally, OCA’s blogger responds. In part:

Dan Clancy rebuts our post by describing the benefits the Settlement will bring to partner libraries... [W]e remind our readers that libraries only receive [digital] copies if they join the Settlement and enter into a new legal agreement that submits their use of those copies to regulation by the Book Rights Registry.... From the perspective of libraries, the Settlement’s new codification of rights is not an improvement upon the allowances of copyright law. [T]he Settlement [offers] no such promises of a deeply discounted subsidy. [I]t says only that Google “may subsidize the purchases of Institutional Subscriptions by [partner] libraries.” So, this is good news, but it’s not in the Settlement...

Let’s restate the obvious: libraries are not a partner to the Settlement. Only Google, the Authors Guild, and the Association of American Publishers will be signing the document. Although a few libraries participated in the negotiations, their interests did not animate the agreement, as is evident from revelations since the settlement was announced. One participant has written: “Libraries [were] not sitting at the head of the bargaining table, and they [were] not going to be able to get everything they wanted, or perhaps even much of what they wanted.” Furthermore, one of Google’s library partners, Harvard University, has refused to join the Settlement. Even Google’s biggest library supporters (Stanford, Michigan, and the University of California) have admitted that they “have not unanimously agreed to all aspects of the proposed settlement.”...

The OCA seeks to foster the conditions that will allow libraries to flourish, to expand, and to continue to provide the greatest possible free access to human knowledge while at the same time preserving that knowledge. We don’t believe this settlement will nurture that ideal...

## ContentBlogger

John Blossom posted “Book deal Googled: Out-of-print books come out from the snippet-hole” on October 31, 2008 ([www.shore.com/commentary/weblogs](http://www.shore.com/commentary/weblogs)). Excerpts, noting that this is distinctly a publishing-industry source:

[A]t the end of the day most of the several years between Google’s introduction of its book scanning program...and the recently announced settlement with the book industry for USD 125 million has been a matter of the book publishing industry deciding to name a reasonable price that would sync up with the realities of book publishing in an electronic marketplace...

In many ways this enables the book industry to monetize fringe content far more effectively via Google partners such as Amazon, in essence validating the value of Chris Anderson’s “long tail” theory for content that was sometimes discounted by book industry executives resistant to Google’s scanning efforts. The settlement is really just a bulk licensing fee to make it easier to administer long-tail revenues, not too different than the industry royalties paid by radio stations. This sets up people to buy books in print and in e-reading devices like Amazon’s Kindle based on Google Books “broadcasts” just as premium downloads and CDs are fed by online and broadcast radio revenues. With finding an audience for one’s content the greatest challenges for all publishers Google Books has become a powerful browsing engine that maximizes the value of any title, new or old, for an audience that is just right for it...

So all in all this deal is likely to turn into a content industry love-fest over the next few years, a peace treaty that finally enables book publishers to leverage the vast power of Google’s book scanning initiative, thus avoiding expensive or less powerful alternatives and enabling book marketers to accelerate their increasingly aggressive exploitation of online channels for their marketing efforts... :et’s all just be glad that there are better times ahead for book publishers who are learning how to exploit electronic content markets far more effectively.

Blossom seems to assume (in sections not quoted here) that Google *will* make OP items available for download to ebook devices, reading rather a lot into the settlement. His perspective certainly suggests that publishers are the winners, with ordinary citizens having yet one more way to buy things.

## Deutsche Welle

Here’s an odd one—or maybe not. “German publishers accuse Google controlling culture” appeared October 30, 2008 in this English-language German site ([www.dw-world.de/dw/](http://www.dw-world.de/dw/)). Excerpts:

The Boersenverein, the German booksellers and publishers association which has bitterly opposed Google for years, rejected the accord as a “creeping takeover.” “This accord is like a Trojan Horse,” Alexander Skipis, chief executive of the Boersenverein, said in a statement on Thursday, Oct. 30. “Google aims to achieve worldwide control of knowledge and culture. In the name of cultural diversity, this American model is out of the question for Europe,” he said, adding that it contradicted “the European ideal of diversity through competition.”...

The Boersenverein has funded a pay-for-use book-scanning service for German-language books, Libreka....

In the United States the Google accord has been widely welcomed, since the bulk of books existing today are hard to obtain, as they are no longer on sale and uneconomic to reprint though their copyrights have not expired....

Since the settlement *only* affects U.S. users, it’s hard to see the point of this—but there is one interesting note here, in addition to general cultural paranoia: “The Boersenverein has funded a pay-for-use book-scanning service for German-language books, Libreka.”

## James Gibson

This commentary, “Google’s new monopoly?” appeared November 3, 2008 in the *Washington Post*. Gibson directs the Intellectual Property Institute at the University of Richmond School of Law. Excerpts:

...Google seemed like a copyright owner’s worst nightmare: a risk-taking iconoclast with deep pockets, unafraid to litigate licensing issues all the way to the Supreme Court. So the copyright industry held its breath as the controversy played out, wondering if it had met its match.

Viewed in this light, the settlement looks like a setback for Google. In the game of brinksmanship, Google blinked—losing its nerve like so many copyright defendants do. In reality, however, settling probably puts Google in a better position than it would have been if it had won its case in court.

Here’s why: Google’s concession has made it more difficult for anyone to invoke fair use for book searches. The settlement itself is proof that a company can pay licensing fees and still turn a profit. So now no one can convincingly argue that scanning a book requires no license. If Microsoft starts its own book search service and claims fair use, the courts will say, “Hey, Google manages to pay for this sort of thing. What makes you so special?”

By settling the case, Google has made it much more difficult for others to compete with its Book Search service...

## Consilience

Greg Grossmeier, a student in Michigan's School of Information who works for Creative Commons but also for Paul Courant, posted "Google Book settlement" on November 8, 2008 (blog.grossmeier.net). Excerpts from a trailing set of issues, following favorable comments on some good aspects of the settlement:

[T]he fact that this is going to be a "**Universal Bookstore**" not a "Universal Library" is slightly saddening. I don't have a legal reason to feel sad; the copyright holders have every right to charge for these materials. But I feel like everyone other than Google, the authors, and the publishers are being scammed. Again, not for a legal reason, but for a moral reason:

Libraries, through public funding, have been keeping these books safe... These books, up until the day of the settlement, [were] worthless to the publishers and authors... Now, Google, through its Universal Bookstore, will sell you these books and pay the authors for them. Google will not pay the libraries who were the ones who made this whole endeavor possible. Sure, the libraries agreed to only get the digital copies back as part of their agreements with Google, but that was before anyone had thought about this possibility. Should those contracts be renegotiated?

**What Happened to Fair Use?** This could possibly be one of my biggest critiques of this settlement: the pure fact that there is a settlement.... Google had a fairly good Fair Use argument and may have indeed won the case based on it. This would have been a *great thing* (most likely). Others would have the same rights as Google as it pertains to the scanning and displaying of books. Now, however, Google is a "special citizen" in this arena; they have "rights" others do not. Is that fair? No. Is [this] best for our future, and the future of libraries? No.

Hopefully I don't sound too negative towards this settlement. OK, let's be honest, I am pretty darn negative towards it. But hey, that is my job, at least what I see my job being. There are plenty of people out there being paid a large sum of money to tell you how good this settlement is. The ones who are out there telling you how bad it is are most likely not being paid to do so; I'm not.

I think it's a bit unfair (and degrades an otherwise interesting discussion) to assume that most or all of those who think the settlement's a good thing are being "paid a large sum of money" to do so—or, for that matter, that none of those upset with the settlement are being paid to write about it.

## The laboratorium

James Grimmelman (noted earlier) offered "Principles and recommendations for the Google Book

Search settlement" in a massive November 8, 2008 post. How massive? More than 7,000 words, reflecting a careful review of the entire proposed settlement (including appendices) and discussions with "a number of my favorite smart people, some in Google's pocket, some opposed to all things Google."

There's way too much to include here if this article is to fit within a single issue of *C&I*, and it's dense writing in the right way—that is, Grimmelman's saying a lot that's worth thinking about with little excess verbiage. For that reason, I'm including the full URL ([http://laboratorium.net/archive/2008/11/08/principles\\_and\\_recommendations\\_for\\_the\\_google\\_book](http://laboratorium.net/archive/2008/11/08/principles_and_recommendations_for_the_google_book)) and *strongly* recommend that you read it yourself. Here, I'm only quoting portions of the start of the essay, and the summary of principles and recommendations (for modifying the settlement) that ends the essay. Those principles and recommendations are based on and informed by the rest of the essay.

My starting point is that the settlement is a good thing. Everyone is better off than in a world where the alternative is no Google Book Search.

- Google will take in a lot of money selling e-books to consumers, subscription databases to libraries, and book search ads to advertisers.
- Authors and publishers will receive the majority of that money. They can choose the price they sell individual copies at; they'll get a proportion of the revenues from other uses based on how popular their books are.
- Public and nonprofit libraries will get at least some minimal all-you-can-drink privileges at the fire hose.
- Universities, schools, and lots of other institutions will be able to subscribe to the fire hose of books, as well.
- The libraries participating in scanning books will get back digital copies of the books from their collections. While there will be usage restrictions on the in-copyright ones, the digital copies of the public-domain ones are not to be sneezed at.
- The public as individuals get an incredibly useful book search engine, one that will come increasingly close to being genuinely comprehensive over time. We also get another convenient source of e-books, free PDF access to millions upon millions of public-domain books, and some degree of full-text library-based access to the rest.
- The public at large gets a substantial leg up on solving the orphan works problem. This system will encourage some copyright owners to come forward, will enable many sensible uses of many books for which no copyright owner can be



found, and will help in cleaning up the records to help track down copyright owners in general.

These are serious benefits, and the settlement is a universal win compared with the status quo.

After several thousand, here's the summary of principles (P#) and recommendations (R#):

P0: The settlement should be approved

R0: Approve the settlement.

P1: The Registry poses an antitrust problem

R1: Put library and reader representatives on the Registry's board.

R2: Require the Registry to sign an antitrust consent decree.

R3: Give future authors and publishers the same deal as current ones.

P2 If it didn't already, Google poses an antitrust problem

R4: Strike the most-favored-nations clause.

R5: Allow Google's competitors to offer the same services the settlement allows Google to offer, with the same obligations.

R6: Authorize the Registry to negotiate on copyright owners' behalf with Google's competitors.

P3: Enforce reasonable consumer-protection standards

R7: Prohibit Google from price discriminating in individual book sales.

R8: Insert strict guarantees of reader privacy.

R9: Protect readers from being asked to waive their rights as a condition of access.

P4: Make the public goods generated by the project truly public

R10: Require that Google's database of in-print/out-of-print information be made public.

R11: Require that the Registry's database of copyright owner information be made public.

R12: Require the use of standard APIs, open data formats, and (for metadata) unrestricted access.

P5: Require accountability and transparency

R13: Require that Google inform the public when it excludes a book for editorial reasons.

R14: Tighten up the definition of "non-editorial reasons" for excluding a book.

R15: Allow any institution ready, willing, and able to participate in scanning books to do so.

A few of those may not make sense out of context. R8 and R9 seem particularly important from a library viewpoint (and R1 wouldn't hurt), but the essay makes cogent cases for all of them. Read the comments as well.

### **@collectanea**

Georgia Harper, then a virtual scholar at the Center for Intellectual Property (University of Maryland Uni-

versity College) wrote several posts on the proposed settlement ([chaucer.umuc.edu/blogcip/collectanea/](http://chaucer.umuc.edu/blogcip/collectanea/)). I think it's fair to characterize her views as pro-Google, making the most charitable interpretations of Google's intentions when they're open to question. In "Google Book Search and orphan works" (posted November 1, 2008), Harper discusses the potential impact on orphan works. Excerpts:

...This is the publisher's and Google's no nonsense business approach: "Hey, let's just start selling all the books and if there's money to be made, the owners will either show up to claim it, or the money will lie there for 5 years while we give everyone time to wake up and smell the coffee. At the end of 5 years, we'll pretty much know what's orphan and what's not. What's not to like?"

At first I was appalled. Especially because the settlement terms provided that the information about who claimed what was going to be kept secret between Google and the publishers/authors (ie, the Registry)...

...I'm happy that in five years...there will (we take on faith) be some sort of way to pull together which books have not been claimed and more or less know what's orphaned of those works that were published in the 20th century. But the process by which a book is claimed needs to be transparent. If the public will not know whether claimants meet rigorous or absurdly simple criteria for proving their claims, confidence in the outcome of the process will fail. This has the potential to be very powerful—or a joke... Imagine if the process of registering a copyright at the Copyright Office were secret and only the result, that a copyright was registered, were available. No actual registration, no basis for disputing whether a claim is valid...

I want this process to work. I think it has a much better chance of working than that piece of, uh, than that piece of [orphan works] legislation that nearly passed earlier this fall. It doesn't give us an answer today and it *only* deals with books, so it's not a comprehensive solution, but it might serve as an example of what works, assuming it does work. But libraries can still do their own research on individual titles that they think may be orphans while we wait for this deal's market incentives to do their job, and for it to become clear that transparency is in the owners' best interests as well as the public's...

Speculation is fun. But this deal offers a real living, breathing experiment for bringing orphan works to a new audience, and for bringing information about what works are orphans to light as well. The settlement is not written in stone. I know from working with Google as a Book Search Partner that Google doesn't work at the level of its contractual commitments. It sees those commitments as starting points and works up from there. If there are aspects of the

settlement that threaten its value, they will be addressed. I think the transparency of the Registry process and outcomes is one of those elements.

Will the settlement improve the possibility of reforming use of orphan works? Only time will tell.

### *Librarian on the edge*

It's possible to read the proposed settlement from a library perspective and be unabashedly enthusiastic. Here's Terry Ballard on November 18, 2008, in a post titled "Surprise! Google just gave your library millions of books for free" ([librariansonedge.blogspot.com](http://librariansonedge.blogspot.com)):

...The news is better than anything I could have imagined, and I've got a pretty good imagination. Here are some of the headlines.

The biggest shift in access comes with the books that are still in copyright but out of print. Now, although they are fully indexed, you can only see snippets of text. After the new plan goes into effect, you can see 5 page blocks of text, up to 20% of the book. Considering that this accounts for 70% of their online collection, this is a massive increase in the book data available online...

Books in copyright but out of print will be available for sale by Google, giving the buyer lifetime electronic access to that title. The pricing for most of these books is ten dollars or less.

Books in the public domain will continue to be made available to all users...

Then it gets really interesting. Any public or academic library that requests it can apply to make one terminal a special machine that can have full access to the universe of books that are in copyright but out of print. From the numbers I've heard thrown around, this means that your library just increased by about five million books, and you don't even need to buy new shelves.

One terminal. One terminal on which to read non-downloadable books (and print pages, for a fee). Or, as a comment from Jeff Scott suggests, one terminal to make it clear to your library that it really needs to pay an annual fee for full access to the database—a fee that's based on FTE. How much? Not known. For a library with high demand and budgetary problems, it might be a boon, or it might be a budget-wrecker.

## **Peter Brantley**

Peter Brantley has offered several commentaries on the settlement at his blog ([blogs.lib.berkeley.edu/shime-nawa.php/](http://blogs.lib.berkeley.edu/shime-nawa.php/))—four pages' worth as of early February 2009. Excluding posts that primarily quote others (otherwise quoted here) and arguments already covered elsewhere, Brantley's thoughts add another impor-

tant set of perspectives (even if Brantley does find it necessary to label us old folk as "less digital"). I'm only noting a few of them.

### *An ever sliding window of access*

Excerpts from this October 31, 2008 post:

There is a lot of understandable speculation about the value of the content that has been library-sourced in the Google Book Search settlement proposal. Besides the public domain, lot of the volume is rendered "Not Commercially Available" by definition of the proposal, and therefore uniquely, or nearly uniquely, available through Google Book Search. Libraries are delighted this content can now be made available digitally. And that's a great thing, I agree.

However, in terms of revenue generation, I'd trade a whole backlist for the frontlist: e.g., for the 37 percent (Google's share) of the sales of *Asian Adventures: Hot Nights* at \$7.99 for consumer access. Attention spans are short in the human species, and the transition to digital does benefit long tail reading, but it also arguably biases against older texts. (There are at least two reasons for that bias: older texts are best known by older people, who are less digital; and older books with pre-modern fonts render less successful OCR through high-volume digitization.)

The settlement basis on library-held content assures steady institutional income from libraries etc., at relatively low maintenance costs (heavily front-end subsidized by the participating libraries, who then tax themselves in perpetuity as a class for the opportunity).

That income source aside, the revenue window might be characterized as two fold, direct and indirect. Direct, moving forward, with in-copyright, in-print (commercially available) that can be digitally distributed through various means, and the secondary marketing options therein available...

Indirect benefits come from the registry infrastructure and how that plays out...

And, priceless: knowing through Google IDs who is reading *Asian Adventures*, and other books they are browsing, and what they are hitting on in web searches, and what they are reading in Google News; subscribing to in RSS feeds; what they are looking up in Google Maps; etc. ... That is all pretty nice fodder for advertising and ancillary opportunities.

Notably, only aggregated GBS use-data will be available to the settlement's rightsholders (authors and publishers); none of it is likely to ever be available to institutional or consumer subscribers.

### *Losing what we don't see: Translation*

This November 2, 2008 post notes the difficulty of seeing "what might have been" when looking at a complex text such as the Settlement. In this case, he

notes the extreme conservatism of the Settlement regarding integrity of text (what Google can and can't do to the text). After quoting several paragraphs of the Settlement and noting reasons for its conservatism, Brantley continues:

Without rants or wails, the constraint probably has a lot less to do with the conservatism of publishers, *per se*, than the complicated rights issues that might devolve from derivatives.

And that is where we have to wonder in part what might be lost if some of these battles do not wind up in court. Class action judgments may move us forward in some important ways, but they also close off other paths that are at least as significant in terms of innovation...

Brantley then quotes Ethan Zuckerman on the polyglot internet and the need for “tools and systems to bridge and translate between the hundreds of languages represented online” and notes:

With Google Book Search and the settlement, we have one of the most amazing possibilities: translation of a great corpus of the world's literature. It would not be perfect, but it would be liberating beyond anything we could have presently imagined for delivering information and knowledge into the hands of others. It would not be beyond the pale, e.g., for Google to machine translate all non-fiction works into a limited number of languages, and enable search and translated reading against them. It would not need to assume a high fidelity reconstruction of the text in any language, but a rough and ready translation, presented perhaps through something akin to today's “Snippet view” that preserved the market for authorized translations. Mass machine translation is not a translation of a work, *per se*, but rather a liberation of the constraints of language in the discovery of knowledge.

And it is likely not to be. If that prognosis is true, that is a tragically lost opportunity.

### *Waking up to books in Richmond*

Brantley lives in Richmond, California—one of the less upscale communities in the San Francisco Bay Area. This November 4, 2008 post considers the “free public library access, but only one dedicated terminal per library building” clause. Extensive excerpts:

One of the irksome characteristics of the proposed Google Book Search settlement is the restricted access to the service at public libraries. Public libraries, we must recall, have long been public temples dedicated to equal access; that spirit is enshrined famously at the Boston Public Library—“FREE TO ALL”...

I do not know where program management at Google wakes up every morning; I do not know what pretty suburbs publishing executives wake up in every

morning. But I wake up every morning in the city of Richmond, CA. Richmond is a great city; a city famous for helping win the Second World War; it is in places a beautiful city, and it is a city with incredible promise. It is also a city of underprivileged populations. The reasons for this particular social geography are many, and deeply embedded in historical contexts. But in Richmond, and in many cities around the country, it is heinous to suppose that one public terminal given free reign to the corpus of the world's literature is an adequate set aside against the promise of the opportunity that Google, publishers, and authors have made possible.

Let the population of Greenwich CT and Los Altos CA have their single terminal per library building; they may, by and large, retreat to their homes with high speed internet access, and their schools may very well wind up acquiring their own subscriptions to GBS.

But there will be no deleterious market impact of an expansion of free public access if it was offered in Richmond. Many of my city's fellow residents have *no* internet access at home, *no* (or exceedingly limited) internet access at school. You—Google, publishers, authors—have an incredible opportunity to facilitate learning, reading, questioning.

This is not an economic matter; it is a social foundation. A library is a refuge; you can provide solace in that refuge, and a promise for a different and better kind of future. It is morally incumbent upon you to do so.

I propose that public terminals be accessible on a tiered basis. If a certain percentage of a public library's served population falls beneath the poverty level or a similar metric, the number of public access terminals is commensurately increased.

At public libraries, internet access is a priority; so is access to information. Help them fulfill that promise to those most in need.

I find it hard to regard the “one free terminal per building” clause as being much more than a ploy to encourage public library subscriptions to the licensed database—after all, we're talking about one single physical device at which people would be *reading books*. Not *finding* them: Google Book Search does just as well there. Brantley's points here are well made.

### *Settle for profit or distribution*

On November 11, 2008, Brantley considered the motives of authors and publishers—and some interesting possibilities. Excerpts:

The proposed Google Book Settlement contains a significant number of assumptions concerning the motives of the participants. One of the most obvious

is that authors and publishers are in it, to put it crassly, for the money.

That might not always be a good assumption. Let's say the author of a book, in full possession of rights, wanted to include her book in Google Book Search, and make it freely available for use under a Creative Commons Attribution, NonCommercial, Share-Alike license. The author might not feel the work is likely to generate much additional revenue, or regardless, wants her work to be broadly and widely accessible to the greatest number of readers possible.

How would she do this under the settlement? Is it possible? Would she have to opt out?

The default assumption is that works would be fixed with a pre-determined (algorithmically derived) price... [but that an author can specify their own price]...

Does this mean that the author can set a price of "free"—and more importantly, if they can, is there any mechanism by which to convey the license (Creative Commons or other) under which that freedom of access is governed?

One potential problem with this settlement is its enforcement of a set of license and pricing assumptions that might be difficult to unbundle. In this fashion, the settlement assumes the primary motivation is profit maximization, not distribution to the greatest number.

The first comment notes that a rightsholder could always offer CC rights elsewhere—but that doesn't help if the rightsholder doesn't have a digital version (likely the case for older books). The second comment, from Dan Clancy of Google, is clear enough:

Good question Peter. The simple answer is yes. A rightsholder will be able to set the price to 0 or for that matter any other price they might desire. Similarly there should not be any problem with identifying license terms that they may be using such as Creative Commons. (I am on the book search team and was very involved in the agreement.)

I've seen a statement elsewhere (which may be erroneous) that an author who isn't in the Authors Guild will need to spend \$200 to sign up for the Registry. I own full rights to nine OP books. Five of them have apparently been digitized already; I assume the other four will show up sooner or later. I'd be *delighted* to make them available for \$0 under a Creative Commons BY-NC license. Am I willing to spend \$200 in order to give away Google-scanned versions of these books? Not really. Still, this is a very good provision.

## Karen Coyle

Karen Coyle has posted a *lot* about the settlement at *Coyle's InFormation* ([kcoyle.blogspot.com](http://kcoyle.blogspot.com))—I count at

least eleven posts through February 11, 2009, and I may be missing some of them. The label "google-books" will find most (but not all) of them. Just a few excerpts from some of the posts, where the perspective or material hasn't been covered elsewhere—and it's fair to say that Karen Coyle is not entirely enchanted with the settlement.

### *Google/AAP settlement*

This November 3, 2008 post gets things off to a lively start, with a post that's hard not to quote in its entirety. But I'll try...

This Google/AAP settlement has hit my brain like a steel ball in a pinball machine, careening around and setting off bells and lights in all directions. In other words, where do I start?

Reading the FAQ...it seems to go like this:

Google makes a copy of a book.

Google lets people search on words in the book.

Google lets people pay to see the book, perhaps buy the book, with some money going to the rights holder.

Google manages all of this with a registry of rights.

Now, replace the word "Google" above with "Kinko's."

Next, replace the word "Google" above with "A library."

*TILT!* If Google is allowed to do this, shouldn't anyone be allowed to do it? Is Jeff Bezos kicking himself right now for playing by the rules?...

**Ping!** Next thought: we already have vendors of e-books who provide this service for libraries. They serve up digital, encoded versions of the books, not scans of pages... The current Google Books offering is very feature poor. Also, because it is based on scans, there is no flowing of pages to fit the screen. The OCR is too poor to be useful to the sight-impaired...

**Ping!** [Quotes the one-free-terminal library clause] *TILT!* Were any public libraries asked about this? Does anyone have an idea of what it will cost them to 1) manage this limited access and pay-per-page printing 2) obtain more licenses when demand rises? Remember when public libraries only had one machine hooked up to the Internet? Is this the free taste that leads to the Google Books habit?

**Ping!** The e-book vendors only provide books where they have an agreement with the publishers, thus no orphan works are included. So, will Google's niche mainly consist of providing access to orphan works? Or will the current e-book vendors be forced out of the market because Google's total base is larger, even though the product may be inferior?...

*TILT!* Rights holders can opt-out of the Google Books database. If (when) Google has the monopoly on books online, opt-out will be a nifty form of censor-

ship. Actually, censorship aimed directly at Google will be a nifty form of censorship.

*GAME OVER. All your book belong to us.*

I question the use of “censorship” to describe the act of an author choosing not to make an OP book more widely available than it currently is. One statement is simply wrong: Google won’t manage the Registry. Otherwise, some interesting questions.

### *Google giveth...and taketh away*

This November 18, 2008 post should clarify Coyle’s stance, if the post above seems ambiguous. I’m omitting material covered elsewhere. A few excerpts from a long post, with a few interleaved notes:

The agreement does not answer the all-important question of whether scanning for the purposes of searching is an allowed use under copyright law.

**Nor could it:** It’s a settlement, not a judicial finding.

The agreement flaunts the concept of Fair Use by quantifying the amount of an in-copyright book that users can view for free (“20% of the text,” “five adjacent pages,” but not the final 5% of a fiction book, to keep the endings a surprise.) The ARL document has Google saying that it will not interfere with fair use. I can’t find that statement in the actual settlement. These quantities are contractual, and I’m assuming that that technology will not allow users to exert fair use rights, only the contractual agreement.

I disagree that this has anything *at all* to do with fair use. It’s a set of provisions for online access and has *nothing* to do with your rights to use all or part of a published or purchased work. I can think of no plausible interpretation of fair use that would require a publisher to *facilitate* digital copying of something that’s not originally in digital form.

### **Key Points Relating to Libraries**

This is the hard part for me. Hard in that it really hurts.

- After digitizing books held in libraries, Google will then turn around and become a library vendor, supplying those same books back to libraries under Google’s control. Each public library in the US will get a single “terminal” provided (and presumably controlled) by Google...
- Libraries and institutions can also subscribe to all or part of the database of out of print books. Access is not perpetual, but limited to the life of the subscription.
- There is verbiage about how users in these institutions can share their “annotations.” In other words, if you take notes on your own, obviously those are yours. But if you use the capabilities of the system to make your notes in the system, you cannot share your own notes freely.

### **Now for the Clincher**

... this is the pact with the devil.

- A library can partner with Google for digitization of its collection and get the same release from liability that Google has. The library can keep copies of these digitized books, however, it must follow security standards set by Google and the AAP...
- Libraries that make this pact with the devil are thereby allowed to preserve the files, print replacement copies for deteriorating books, and provide access for people with disabilities. Note that all of these uses by libraries are *already* allowed by copyright law.
- The libraries that make this pact with the devil *cannot* let their users read the digitized books. Well, they can let them read up to five (5!) pages in any digitized book. Presumably if the library wants to provide other uses it must subscribe to Google’s service....

... and if you refuse to negotiate with the devil...

- Current Google library partners who do not choose to become party to this must delete all copies of digitizations of in-copyright works made by the Google project in order to obtain a release from liability...
- Even if the library was only allowing Google to digitize public domain works, those libraries must destroy all of their copies to get release from liability in case they misjudged the copyright status of one of those books.

In other words, this agreement is making the assumption that if anyone sues Google for copyright infringement, the library will be a party to that suit.

They say that “the devil is in the details.” In this case that is not true: the devil is right up front, in the main message. That message is that Google has agreed with the publishers, and is selling out the libraries that it has been working with... Participating with Google has been an expensive proposition for the libraries in terms of their own staff time and in the development of digital storage facilities. Part of the appeal of working with Google was the assumption that partnering with the search giant gave the entire project clout and provided some protection for the libraries. With Google and the AAP now in cahoots, the libraries must join them or try to stand alone in an unclear legal situation; an unclear situation that Google invited the libraries into in the first place.

This is classic bait and switch. And it is bait and switch with powerful commercial interests against public institutions. There is no question about it...

### **THIS IS EVIL**

(Orthography, bolding and centering in the original.)  
Evil and the devil: Not a lot of ambiguity there. One

could note (and jrochkind does, in the first comment), that the restrictions on use of digital copies are only on the digital copies *supplied by Google* and that the agreement *specifically* says it does not limit fair use rights (which you can theoretically do in a contractual agreement). One could argue that “In other words” is making huge assumptions.

### *More on Google/AAP*

Posted November 22, 2008. Extracts that aren't covered elsewhere:

#### **Library Involvement**

Some librarians were involved in the settlement talks [working under non-disclosure agreements]... I have heard statements from others who I believe were privy to the negotiations, and they all seem to feel that the outcome was better for libraries due to the involvement of members of our “class.”... Unfortunately that doesn't change my mind about the bait and switch move.

#### **Google Books as Library**

Some have begun to refer to Google Books as a library. We have to do some serious thinking about what the Google Book database really is. To begin with, it's not a research collection, at least not at this point. It's really a somewhat odd, almost random bunch of book “stuff.” As you know, neither Google nor the libraries are selecting particular books for digitization. This is a “mass digitization” project that starts at one end of a library and plows through blindly to the other end. Some libraries have limited Google to public domain works, so in terms of any area of study there is an artificial cut-off of knowledge....

But *most* libraries aren't research libraries—and seven million books counts as a pretty sizable bunch of “book stuff.” Some big libraries *haven't* limited Google to public domain works, so that cutoff simply doesn't exist—otherwise, there would be no *need* for a settlement, as there could be no lawsuit.

So the main reason why Google Books is not a library is that it isn't what we would call a “collection.” The books have not been chosen to support a particular discipline or research area... One of the big gaps in Google Books will be current materials, those that are still in print. Google will need to convince the publishers that it can increase their revenue stream for current books in order to get them to participate.

I don't know of *any* big academic library or public library that's a single disciplinary collection—or, realistically, a set of well-curated collections. As for current materials, Google's doing pretty well with its Google Publisher program.

#### **Subscribing to Google Books: Just Say No?**

Beyond the (undoubtedly hard-won by library representatives) single terminal access in each public library in the US, libraries will be asked to subscribe to the Google Book service in order to give their users access to the text of the books (not just the search capability). This is one of the more painful aspects of the agreement because it seems to ignore the public costs that went in to the purchase, organization, and storage of those works by libraries... The parallels with the OCLC mess are ironic: libraries paying for access to their own materials. So, couldn't the libraries just refuse to subscribe? Not really. Publicly funded libraries have a mission to provide access to the world's intellectual output in a way that best serves their users. When something new comes along—films on DVD, music on CD, the Internet—libraries must do what they can to make sure that their users are not informationally underprivileged. Google now has the largest body of digitized full text, and there will be a kind of “information arms race” as institutions work to make sure that their users can compete using these new resources.

I don't remember public universities admitting to substantial costs in cooperating with Google. It would be interesting to hear such costs. Since *full* access to all OP books, for everyone, for free is *far* beyond any reasonable claims of fair use, I can't imagine that AAP or AG would have ever agreed to such a possibility, and they would surely have had the law on their side.

#### **The (Somewhat Hidden) Carrot**

I can't imagine that anyone thought that libraries and Google were digitizing books primarily so that people could read what are essentially photographs of book pages on a computer screen. Google initially stated that they were only interested in searching the full text of books. While interesting in itself, keyword searching of rather poor OCR text is not a killer app. What we gain by having a large number of digitized books is a large corpus on which we can do computational research...

I have suspected for a while that Google was already doing research on the digital files that it holds. It only makes sense. For academics in areas like statistics, computer science, and linguistics, this corpus opens up a whole range of possibilities for research; and research means grants, and grants mean jobs (or tenure, as the case may be). This will be a strong motivation for institutions to want to participate in the Google Book product...

There's at least one other carrot: Google's said that participating institutions will receive substantial discounts on the subscriptions, possibly discounting to \$0. And, as noted in the comments, the research corpus won't be limited to participating libraries.

## Google and fair use

Excerpts from this December 4, 2008 post, with comments as needed.

...Google's first business is that of indexing resources that are on the web. I'll talk about them as if they were all texts because it's easier, but the same thing could be said for images and other resources.

To do the indexing, Google must make a copy of the web page or document. Using this copy, it adds the page to its search engine....

This is all fine and unremarkable until you look at it from the point of view of copyright law. Copyright is specifically about...making copies, and it gives the right to make copies, or to authorize the making of copies, to the copyright holder. That can be the author, or someone to whom the author has passed along the right...So the big question is: Is Google violating copyright law by making copies of web pages without the permission of the copyright holders? There are two main ways of looking at this:

1. The web is different from the print environment. Anyone who has put their works out on the web has agreed to copying because no one can even view the work without making a copy...

2. The web is **not** different from the print environment. But Google is just producing an index and there is nothing in copyright law that would prevent someone from producing an index of words in texts. The incidental copies that Google makes in order to produce the index are allowed under the Fair Use aspects of the copyright law.

So then we move on to the Google Books project. Initially, Google claimed that it was doing the same thing with books as it does with the web: making incidental copies in order to create keyword indexes to the texts...

In fact, Google did and does make the fair use argument. The libraries that partnered with Google also came to the fair use conclusion in at least some cases...

What was at stake with the AAP lawsuit was exactly this decision about fair use... Although Google has always provided a confident posture to the public, declaring unwaveringly that what it does as a search engine is perfectly within copyright law, the idea of going to court over the issue would have put their entire operation at risk.

Now back to libraries. Fair use is not a list of things you can do but a judgment call relating to some complex factors... [And libraries have somewhat more latitude.] What happened with Google Book Search and the AAP is that the digitization of the libraries' books and subsequent use of those was judged not by the criteria that would be used normally for libraries, of course, but by the criteria that

would be used for a commercial entity. That's totally logical, since although Google was partnered with the libraries, the primary use of the materials was to fuel Google Book Search, an obviously for-profit activity.

Libraries have gotten the short end of the stick because their use of their own materials became commercialized through their partnership with Google. If instead libraries had managed to digitize the books on their own, the outcome would have likely have been entirely different (if any lawsuit had been brought, which might not have happened). I believe that libraries could be found to have a fair use case for digitizing their works for the purposes of searching, and could be allowed to use those digitized copies for the exceptions spelled out in section 108 of the copyright law.... Unfortunately, the concept of digitization of the contents of libraries has now been tainted with the air of commercialization and has earned the wrath of the publishers and authors. The Google/AAP settlement has created a mechanism that ignores the inherent rights of the libraries, but also makes it more difficult for them to justify undertaking their own digitization project...

The settlement might look good from the point of view of a commercial entity facing copyright law, but it binds the non-profit educational and cultural heritage community to legal decisions designed for the for-profit sector. This is not only not a win for libraries, but it will hinder libraries in their efforts to make use of current technologies to further the arts and sciences.

Well...this is not a judicial finding. I find it unfortunate that Google didn't fight the good fight, and I think it will make things *much* harder for another commercial entity to attempt similar digitization and use—but I don't see that library use of "their own materials" has changed in any way.

## Google's gift of books

December 30, 2008—another *long* post. It's about the one-free-terminal provision. Excerpts:

...I'm sure that many folks are quite impressed at this generosity: free access to the public! What's not to like? Well, keep reading.

[I'm omitting a lengthy discussion of the Gates Foundation's library support program, erroneously called Microsoft donations. Coyle says nothing in the description "should be construed to demean the gift from Microsoft" (which wasn't from Microsoft) but it sure doesn't read that way.]

### The First One is Always Free

...Should the single access to the Google Books Public Access Service not suffice, libraries will need to add more subscriptions to meet the demand. It isn't known what this will cost, but unless it is ridiculous-

ly cheap, it eats into the already strained budgets of the libraries. Eventually, the cost will be absorbed into the budget as part of normal expenses, but there will be a painful phase at the beginning. Before they introduce this free service, libraries need to know what the costs will be in 2, 3 or possibly 5 years so they can begin the budget planning process that will allow them to provide full service to their users, if that's what they wish to do...

True, and a major issue. [I'm omitting the "just say no" section, to some extent already covered.]

### **Equal Access for All**

One option that libraries must consider when new services arise that are outside of their budget capabilities is whether they will choose to provide the service with a user fee attached... The public library mission of equal access to all...argues against requiring fees for services, other than those nominal fees designed to prevent squandering of resources (e.g. 25 cents for each book put on hold), or cost recovery for consumable materials, like photocopy services....

We do know that libraries will not be able to offer remote access to their free subscription, only on-site access. That, of course, excludes many users. We also know that there may be advertising included in the service, and it may include the ability to purchase books (online or in hard copy) and additional services. In other words, the library's users become the service's customers...

Story hours exclude users who won't go to the library. In most libraries, book collections exclude those who won't go to the library. Magazines and newspapers in libraries include advertising (and few libraries prohibit the use of Google and other ad-supported websites).

Charity is giving people what they need, not what you want them to have or what you would like them to buy in the future. While the provision of a free, one-user license to libraries may be generous, it is not charitable. It should be viewed in the same way that free samples of cereal are. Actually, the better analogy harks back to the days when cigarette companies gave away free packs of cigarettes on city streets, hoping to encourage non-smokers to become smokers...

Since I don't remember a claim of charity, attacking this as not being charity is pointless—but I tend to agree that it falls into "the first one is free" category, associated as much with (illegal) dope peddlers as with cigarette companies.

### ***Google Books and social responsibility***

This post appeared on January 10, 2009. I take issue with the very first sentence, as I've taken issue consistently with the same claim by others with even higher profiles than Coyle (who are even less likely to ever

admit they could be mistaken). Extensive excerpts and comments:

The digitization of books by Google is a massive project that will result in the privatization of a public good: the contents of libraries. While the libraries will still be there, Google will have a de facto monopoly on the online version of their contents.

**Nonsense.** Sheer, utter nonsense. The libraries and contents *will* still be there. OCA *will* still be there. I'm sorry, but this one just drives me nuts: It's demonization of the worst kind and an abuse of the language.

While regulation of industry has fallen out of favor in these 'free market' times, we do have a history of making particular demands on companies whose products and services have an important social impact, such as broadcast television or telephone services...

If I were in a position to require social responsibility of Google and its digitization program, these would be my terms:

### **Sustainability**

While Google is a hot company today, it may not last forever. Actually, it probably won't be around for the 200-odd years that have been covered by the libraries it is working with. To protect against the loss of the digitized books should Google either disband or decide not to continue the Books product line, Google should be required to place the digital copies in escrow, where they will be preserved. My preference would be for the escrow body to be a public institution (or a group of such institutions) that has proven longevity and stable public support.

Won't the fully-participating libraries have digital copies? I can't think of institutions with better longevity.

### **Intellectual Freedom**

The First Amendment prevents the government from censoring its citizens, and we rely heavily on this key right as the basis for many of our freedoms. Private companies are not bound by the First Amendment; as a matter of fact, in law they are protected by it as honorary persons. This means two things: first, that private companies can (and do) censor their products, and second, that they can be held liable for any social harm that is perceived if they do not censor. Thus publishers can be held liable for errors of fact in the books they produce, or a company that promises a 'child friendly' web site can be held liable if pornography slips through their filter.

**TILT!** (To use Coyle's mannerism.) That ain't censorship. That's the decision not to publish—without which the First Amendment becomes meaningless.

I want Google to have the same right to deliver books to users that publicly funded libraries do. How this could be worked out in terms of law and liability I



must leave to others to determine, but what I am thinking of like the of common carrier model that has been used for communications companies. Basically, Google should be required to carry all digital Books without discrimination and without liability.

You mean “all digital books that Google’s scanned”? I suspect Google wouldn’t argue with this.

### Privacy

Public libraries are bound by state laws to protect the privacy of their users. This protection generally takes the form of enforced confidentiality over any records of library use. This is, in a sense, the other side of the intellectual freedom coin: people are only free to access the speech of others if they are guaranteed that they will not be watched or tracked, and that their information access will not be revealed to others. There are no laws that bind private companies to this same standard, but companies are held to their own promises of privacy to their users. Google should develop a particularly strict privacy policy for the Books product, and should be willing to allow auditing of its practices so that users can trust the company’s practices. Libraries themselves will insist on such a guarantee if they are to include the Book product in the services they provide to their own users.

**Absolutely** agreed (although not all states have confidentiality laws, and the Federal government will, demonstrably, ignore such laws). I like the “will” in the last sentence—unfortunately, it should be a “should,” since far too many librarians adopt a “who really cares?” attitude about confidentiality.

### Transparency

...If the Book product will be licensed by educational institutions, it has to be possible for those institutions to know the status of works and to understand what decisions can be made. Transparency also implies a process for appeal or at least discussion with the vendor about decisions, because those decisions will affect the value the product has in our environments....

Certainly desirable. Frankly, if it wasn’t for the oft-repeated nonsense in the first paragraph, I would have little trouble with this post in general.

I’ll skip a post from an ALA Midwinter panel, except for one paragraph that’s not all that surprising:

Google itself is not thrilled about becoming a library vendor, because it recognizes that it’s not a big bucks market and it doesn’t fit into the Google business model well. (At one point Dan mentioned that getting checks for \$5000 from public libraries isn’t very appealing.)

It’s not a *small* bucks market—some STM publishers and aggregators are doing pretty well—but it *doesn’t* fit into the Google business model. That’s one reason I

question the demonization of Google on this point: I suspect this was the best deal they could negotiate.

A January 26, 2009 post raises 36 questions. Coyle says “we will try to find some definite place to put these”—and I’d suggest ALA OITP is the proper place. A few of the questions have already been answered, and partial answers to some of the others are already extant—but most of them just won’t be answered for a while.

Finally, a January 28, 2009 post offers a version of Coyle’s talk during a Google panel at ALA Midwinter. (Midwinter supposedly doesn’t have programs other than the ALA President’s Program. Sometimes a tough distinction to make—as is also true for the LL-TA Top Tech Trends, which Coyle was also part of.)

Overall: I believe Coyle demonizes Google too readily and abuses common sense in some areas—but she also raises important questions.

## Robert Darnton and Paul Courant

Harvard’s Robert Darnton was the most outspoken critic of the settlement among library directors whose libraries are part of the Google Library Project. Probably the best available version of Darnton’s thoughts on the topic appear in the February 12, 2009 *New York Review of Books*, in the article “Google & the future of books” ([www.nybooks.com/articles/22281](http://www.nybooks.com/articles/22281)). Much of the essay deals with broader matters, and I’m omitting most of that—but suggest that you should read it. Some excerpts from the later portions, dealing more directly with Google:

...When businesses like Google look at libraries, they do not merely see temples of learning. They see potential assets or what they call “content,” ready to be mined. Built up over centuries at an enormous expenditure of money and labor, library collections can be digitized en masse at relatively little cost—millions of dollars, certainly, but little compared to the investment that went into them.

Libraries exist to promote a public good: “the encouragement of learning,” learning “Free To All.” Businesses exist in order to make money for their shareholders—and a good thing, too, for the public good depends on a profitable economy. Yet if we permit the commercialization of the content of our libraries, there is no getting around a fundamental contradiction. To digitize collections and sell the product in ways that fail to guarantee wide access would be to repeat the mistake that was made when publishers exploited the market for scholarly journals, but on a much greater scale, for it would turn the Internet into an instrument for privatizing knowledge that belongs in the public sphere...

...Four years ago, Google began digitizing books from research libraries, providing full-text searching and making books in the public domain available on the Internet at no cost to the viewer... Everyone profited, including Google, which collected revenue from some discreet advertising attached to the service, Google Book Search. Google also digitized an ever-increasing number of library books that were protected by copyright in order to provide search services that displayed small snippets of the text...

[Summary description of the suit and the proposed institutional license, one-terminal access license and consumer license.]

After reading the settlement and letting its terms sink in—no easy task, as it runs to 134 pages and 15 appendices of legalese—one is likely to be dumbfounded: here is a proposal that could result in the world's largest library. It would, to be sure, be a digital library, but it could dwarf the Library of Congress and all the national libraries of Europe. Moreover, in pursuing the terms of the settlement with the authors and publishers, Google could also become the world's largest book business—not a chain of stores but an electronic supply service that could out-Amazon Amazon...

Who could not be moved by the prospect of bringing virtually all the books from America's greatest research libraries within the reach of all Americans, and perhaps eventually to everyone in the world with access to the Internet?...

Unfortunately, Google's commitment to provide free access to its database on one terminal in every public library is hedged with restrictions... But Google's generosity will be a boon to the small-town, Carnegie-library readers, who will have access to more books than are currently available in the New York Public Library. Google can make the Enlightenment dream come true.

But will it? The eighteenth-century philosophers saw monopoly as a main obstacle to the diffusion of knowledge—not merely monopolies in general, which stifled trade according to Adam Smith and the Physiocrats, but specific monopolies such as the Stationers' Company in London and the booksellers' guild in Paris, which choked off free trade in books.

Google is not a guild, and it did not set out to create a monopoly. On the contrary, it has pursued a laudable goal: promoting access to information. But the class action character of the settlement makes Google invulnerable to competition... If approved by the court—a process that could take as much as two years—the settlement will give Google control over the digitizing of virtually all books covered by copyright in the United States.

This outcome was not anticipated at the outset. Looking back over the course of digitization from the

1990s, we now can see that we missed a great opportunity. Action by Congress and the Library of Congress or a grand alliance of research libraries supported by a coalition of foundations could have done the job at a feasible cost and designed it in a manner that would have put the public interest first... It is too late now. Not only have we failed to realize that possibility, but, even worse, we are allowing a question of public policy—the control of access to information—to be determined by private lawsuit.

While the public authorities slept, Google took the initiative. It did not seek to settle its affairs in court. It went about its business, scanning books in libraries; and it scanned them so effectively as to arouse the appetite of others for a share in the potential profits...

As an unintended consequence, Google will enjoy what can only be called a monopoly—a monopoly of a new kind, not of railroads or steel but of access to information. Google has no serious competitors...

Google's record suggests that it will not abuse its double-barreled fiscal-legal power. But what will happen if its current leaders sell the company or retire?... What will happen if Google favors profitability over access? Nothing, if I read the terms of the settlement correctly...

Free-market advocates may argue that the market will correct itself. If Google charges too much, customers will cancel their subscriptions, and the price will drop. But there is no direct connection between supply and demand in the mechanism for the institutional licenses envisioned by the settlement. Students, faculty, and patrons of public libraries will not pay for the subscriptions. The payment will come from the libraries; and if the libraries fail to find enough money for the subscription renewals, they may arouse ferocious protests from readers who have become accustomed to Google's service. In the face of the protests, the libraries probably will cut back on other services, including the acquisition of books, just as they did when publishers ratcheted up the price of periodicals.

No one can predict what will happen. We can only read the terms of the settlement and guess about the future. If Google makes available, at a reasonable price, the combined holdings of all the major US libraries, who would not applaud? Would we not prefer a world in which this immense corpus of digitized books is accessible, even at a high price, to one in which it did not exist?

Perhaps, but the settlement creates a fundamental change in the digital world by consolidating power in the hands of one company...

Whether or not I have understood the settlement correctly, its terms are locked together so tightly that they cannot be pried apart... Yet this is also a tipping

point in the development of what we call the information society. If we get the balance wrong at this moment, private interests may outweigh the public good for the foreseeable future, and the Enlightenment dream may be as elusive as ever.

Paul Courant responded, in a letter that appears in full form as a February 4, 2009 post on *Au Courant* (paulcourant.net). Excerpts:

My colleague and friend Robert Darnton is a marvelous historian and an elegant writer. His utopian vision of a digital infrastructure for a new Republic of Letters... makes the spirit soar. But his idea that there was any possibility that Congress and the Library of Congress might have implemented that vision in the 1990s is a utopian fantasy. At the same time, his view of the world that will likely emerge as a result of Google's scanning of copyrighted works is a dystopian fantasy.

The Congress that Darnton imagines providing both money and changes in law that would have made out-of-print but in-copyright works (the great majority of print works published in the 20th century) digitally available on reasonable terms showed no interest in doing anything of the kind... The committees that write copyright law are dominated by representatives who are beholden to Hollywood and other rights holders. Their idea of the Republic of Letters is one in which everyone who ever reads, listens, or views pretty much anything should pay to do so, every time.

The Supreme Court, which was given the opportunity to limit the extension of the term of copyright, which was already far too long... refused to do so...

In short, over the last decade and more, public policy has been consistently worse than useless in helping to make most of the works of the 20th century searchable and usable in digital form. This is the alternative against which we should evaluate Google Book Search and Google's settlement with publishers and authors.

First, we should remember that until Google announced in 2004 that it was going to digitize the collections of a number of the world's largest academic libraries, absolutely no one had a plan for mass digitization at the requisite scale. Well-endowed libraries, including Harvard and the University of Michigan, were embarked on digitization efforts at rates of less than ten thousand volumes per year. Google completely shifted the discussion to tens of thousands of volumes per week, with the result that overnight the impossible goal of digitizing (almost) everything became possible. We tend to think now that mass digitization is easy. Less than five years ago we thought it was impossibly expensive.

The heart of Darnton's dystopian fantasy about the Google settlement follows directly from his view that "Google will enjoy what can only be called a mono-

poly ... of access to information." But Google doesn't have anything like a monopoly over access to information in general, nor to the information in the books that are subject to the terms of the settlement... Google is required to provide the familiar "find it in a library" link for all books offered in the commercial product. That is, if after reading 20 percent of a book a user wants more and finds the price of on-line access to be too high, the reader will be shown a list of libraries that have the book, and can go to one of those libraries or employ inter-library loan. This greatly weakens the market power of Google's product. Indeed, it is much better than the current state affairs, in which users of Google Book Search can read only snippets, not 20% of a book, when deciding whether what they've found is what they seek.

Darnton is also concerned that Google will employ the rapacious pricing strategies used by many publishers of current scientific literature, to the great cost of academic libraries, their universities, and, at least as important, potential users who are simply without access. But the market characteristics of current articles in science and technology are fundamentally different from those of the vast corpus of out-of-print literature that is held in university libraries and that will constitute the bulk of the works that Google will sell for the rights holders under the settlement agreement. The production of current scholarship in the sciences requires reliable and immediate access to the current literature. One cannot publish, nor get grants, without such access. The publishers know it, and they price accordingly. In particular the prices of individual articles are very high, supporting the outrageously expensive site licenses that are paid by universities. In contrast, because there are many ways of getting access to most of the books that Google will sell under the settlement, the consumer price will almost surely be fairly low, which will in turn lead to low prices for the site licenses. Again, "find it in a library," coupled with extensive free preview, could not be more different than the business practices employed by many publishers of scientific, technical and medical journals.

There is another reason to believe that prices will not be "unfair", which is that Google is far more interested in getting people to "google" pretty much everything than it is in making money through direct sales...

The settlement is far from perfect. The American practice of making public policy by private lawsuit is very far from perfect. But in the absence of the settlement—even if Google had prevailed against the suits by the publishers and authors—we would not have the digitized infrastructure to support the 21st century Republic of Letters. We would have indexes and snippets and no way to read any substantial

amount of any of the millions of works at stake on line. The settlement gives us free preview of an enormous amount of content, and the promise of easy access to the rest, thereby greatly advancing the public good.

Of course I would prefer the universal library, but I am pretty happy about the universal bookstore. After all, bookstores are fine places to read books, and then to decide whether to buy them or go to the library to read some more.

With writers as eloquent and well-informed on the issues as Michigan's Paul Courant and Harvard's Robert Darnton, I'm disinclined either to comment or to get in the middle. In this case, I believe both are, to some extent, right—as are, to a great extent, *most* of the apparently contradictory perspectives offered in this roundup.

### Open issues

Among other things, we don't know how long this will take—and, crucially for many libraries, how much the subscription database will cost. We also don't have much of an overall sense of how good (or bad) those scans really are.

Stay tuned.

## The Last Word (So Far)

I'm writing this in February 2009. While it appears that the proposed settlement is on the fast track for approval, that won't happen until at least May 2009—and it's likely to be a while before that, maybe even a year or two, before some of the unanswered questions get answered.

I strongly suspect we will **not** see major changes in the terms of the settlement, and especially not changes that work in the direction most library critics would like. (I'm an optimist by nature and sometimes accused of being a Candide, but I'm not optimistic enough to believe that scenario.)

My own thoughts—as a library person, a believer in fair use, one who is cautious about letting Google take over too much of my life, a book reader, but also as an author with several OP books for which I am the sole rightsholder (at least five of which have been scanned by Google)—appear at the start of the commentary and scattered throughout. I fall squarely into the “mixed feelings” category, with a number of regrets and, frankly, a strong sense that Google's plans to sell access to individual OP books may turn out to be a bad mistake, given the quality of some of the scans I've seen.

But I'm not going to claim the last word—at least not all by myself. On February 4, 2009, Emily Ford posted “My (our) abusive relationship with Google

and what we can do about it” at *In the library with the lead pipe* ([inthelibrarywiththeleadpipe.org](http://inthelibrarywiththeleadpipe.org)). The post—really a short peer-reviewed article, as is customary for this unusual blog—reminds me *why* I'm reluctant to switch from Bloglines to Google Reader, why I'm uninterested in GoogleDocs (there are other reasons, and I really like Word2007), and why I make an effort to “switch up” searches to Yahoo! and Windows Live. Mostly, however, it offers a thoughtful look at the settlement and offers some useful suggestions. Extended, slightly edited excerpts follow (yes, the blog has the same BY-NC Creative Commons license that *C&I* does, so it's legit), with my comments as appropriate. (Since I'm giving Ford's commentary pride of place, it's essential that I be at least as critical of her comments as of anyone else's.)

Since October something has been weighing on my professional mind: my abusive relationship with Google. I love Google, I don't ever want to leave my Gmail, my Gchat, my GoogleDocs, my web searches, my Google Reader, but right now I wish I weren't so dependent on it.

The weight to which I am referring is the proposed Google Book Search Settlement Agreement. Google knows with whom I e-mail and chat, for what I search, what blogs I read, and on and on. With the proposed settlement Google will take a further step in controlling my (and libraries') information use and seeking behavior. Google will know what books I read, what pages I read, how long I read them, what pages I print, and what passages I copy and paste...

One of the comments says Google plans for that *not* to be true, at least for the in-library terminals—but it's certainly a valid concern.

[Summarizes agreement after telling us to read the *2-Page Super Simple Summary*, and says “many of the agreement's facets are antithetical to the mission and purpose of libraries.”]

...What I do want to share is what I think we in the library community can do about the settlement. The stakes of the settlement are enormous, and neither the rightholders nor Google represent libraries in this process. But we, librarians and the library community at large, are an ornery bunch...

Because I don't want libraries, information advocates, patrons, or anyone else to be trapped in an abusive relationship with Google I would like to offer the following suggestions for what individuals and the professional community can do to protect and salvage what remains of our relationship with “the big G.” (And maybe even make this Google Book Settlement Agreement a bit more reasonable.)

### Individuals

### **Educate yourself.**

Knowledge is empowerment. Read through blog posts, documents, and news articles about the proposed settlement agreement. The ALA Washington Office is tracking most everything that's out there and has made a nice little portal web site for you to use. Particularly useful is also the *Guide to the Perplexed: Libraries and the Google Library Project Settlement...*

Because the settlement is so intrinsically tied to copyright law and fair use, this is an ideal time to refresh yourself on the basics. Re-read Kenny Crews' *Copyright Law for Librarians and Educators* and Carrie Russell's *Complete Copyright*. Subscribe to blogs that deal with copyright such as *Copyright advisory network* ([librarycopyright.net/wordpress/](http://librarycopyright.net/wordpress/)) or Karen Coyle's blog.

I'd like to think this extended PERSPECTIVE will also help, although I do point you elsewhere in many cases. I'd also say that you *must* read blogs and other sources critically and perhaps skeptically, certainly but not only Coyle's work.

### **Ruminare.**

Ask yourself and think about the tough questions. During the "Google Book Settlement: What's in it for Libraries?" panel at ALA Midwinter, Karen Coyle posed the following questions: Does the product serve my users? What will the collection be? What is the quality of the product? Panelist Laura Quilter pushed the panel participants and audience to consider the privacy issues presented by the proposed model for accessing digital materials through Google Books. As librarians we have a responsibility to protect our users. Mold and define your personal and professional values for privacy. This will be incredibly useful if you are put in a place to consider purchasing and implementing this subscription product in your library.

### **Be an advocate in your community.**

Let's face it. There are so many issues to follow in our profession, chances are many of your colleagues might not know anything about this proposed settlement agreement. Talk with your colleagues and share with them what you have learned. Push your administrators to find out if any preemptive discussions regarding this product have occurred. What is the institutional stance on the settlement agreement and Google Books in general? By asking the hard questions of our supervisors and administrators, we are often able to generate institutional discourse.

### **The Community**

#### **Ask and discuss.**

ALA has very bright and informed people working to understand the Google Book Settlement agreement. Librarians who specialize in information policy, entire offices and committees that deal with legislation and lobbying for ALA interests. But this 300+ page legal

document...is confusing and still not fully understood by the library community. At the aforementioned Midwinter panel discussion, many things came to light that we (or at least I) did not previously know about the settlement. For example, the settlement will not allow for a subscriber library's users to login via remote access and access their library's subscription to the Google Books database. Users who are community members of a subscribing institution will only be able to access the resource "on campus." Another fine example is how Google will serve public libraries with this product. Google will allow public libraries one access station to the product. Only one.

Not quite true. There will be one *free* terminal per library building (public or academic), and that *free* use must be onsite. Nothing prevents public libraries from subscribing to the database, and it's simply not clear whether subscriptions can provide authenticated remote access.

We need more fora in which to engage to find out exactly what the settlement agreement means to us and our users. Professional organizations, ALA, SLA, PLA, ARL and others should consider hosting more web-hosted seminars for their members on the subject. Moreover, hosting other kinds of discussion fora to ask questions and commiserate within the library community such as BBS or wikis or even blogs will be helpful to those of us who struggle to understand the issues with the settlement.

...Dan Clancy, Engineering Director for the Google Book Search Project, [says] he would like to be able to be available to the library community for more discussion. State libraries, consortia, or other large groups should consider contacting Dan and scheduling teleconferences about concerns.

### **Educate Google.**

I would like to give Google the benefit of the doubt. However, the fact remains that Google is a business and will not implement policy or procedure based upon it being "the right thing to do." Rather, Google will make policy, and change procedure, as it is beneficial to business and the deep Google pocketbook. That being said, I think Google would attempt to take more responsibility for "doing the right thing" if the company were to realize that the proposed settlement model is not one upon which libraries will willingly spend their money. Just because Google will have a monopoly on the digitized books, does not mean we should lower our standards for offering resources to patrons that are easy to use and ethically implemented. We, as a community, need to share with Google the ethical principles and best practices that we have worked so hard to develop—of particular relevance, the *Principles for Digital Content* and the *Principles for a Networked World*.

Two issues here. The minor one: *Principles for Digital Content* has no official standing; it was developed by an OITP working group but has not, as far as I know, been endorsed by the ALA Council. Second, as a commenter says (albeit discussing privacy), there's a real issue of potential hypocrisy here. Do libraries actually hold existing database vendors to these standards? How many libraries have cancelled a full-text serials database because it doesn't meet these principles?

**Develop position statements, draft and pass resolutions, or take other governmental action.**

A unified voice of librarians can be a powerful thing. Moreover, if professional organizations such as ALA [with its 65,000 members] use their position as the good stewards of knowledge and information, we have the ability to put up a good fight that might yield some positive results. Currently the Washington Office is working to gather ALA membership input so that it can issue a position statement or take other action on the settlement...

ALA Council should also consider passing a resolution regarding the Google Book Search Settlement Agreement. It is not out of the question that this kind of political activity will help the organization to retain its integrity and ethics regarding privacy, information policy, and what best serves libraries and patrons.

ALA and other library organizations should consider future legal action. It seems to me that libraries would have a good case to bring forth their own class action lawsuit. This might be a last case resort, but I do not think we should not sit idly by if a large market-driven product were to threaten the library community's ability to best serve the public.

To be honest, I can't imagine the grounds for such a class-action lawsuit or the proposed remedy, since it would be entirely legal for Google to say "OK, then there won't be any full-text view or more than snippets for OP material." What legal grounds would libraries have to challenge such an outcome?

**Create support materials and documents for libraries to use.**

Shortly after the court "okays" the Google Books Settlement agreement, libraries will face a "purchase or not to purchase" question for the Google Books subscription product. Navigating the ins and outs of the legalese in the settlement will be daunting for any library system, consortium, or lone library that chooses to buy the product. Having FAQs handy or even an ALA Toolkit on best implementation practices for Google Books would be a great service.

I'd be astonished if OITP doesn't craft such an FAQ, but it's certainly worth repeating. On the other hand, libraries considering the subscription won't need to

navigate the entire settlement: They'll be offered a contract, which I'm fairly sure won't be 200+ pages long.

**It doesn't have to be a waiting game.**

If we work now to understand what we can about the proposed settlement, if we start to evaluate the effect purchasing this product will have on our libraries and patrons, if we create a unified voice and foster discourse, then we will better be able to keep fires under control and perhaps keep our brains in our heads. Google is a powerful company, but powerful, too, is the voice of libraries and librarians. I firmly believe that if we continue to put our efforts toward understanding everything encompassed by the Google Book settlement issue, then we will better be able to serve our communities, and perhaps inform positive changes that will let us sit in better peace with our friend and enemy. This is my call to you, colleagues, to engage, think, debate, and defend library values. Take control and save yourself from this abusive relationship. Google can be a reference librarian's best friend, but right now, with the proposed settlement, it is looking as if we are subject to continued abuse.

Even though I doubt the agreement will change much, this is an excellent closing statement.

The agreement could be a lot worse. The outcome could also be a lot better. I'm sure Google would agree with both statements, as it finds itself in businesses where it has neither expertise nor much chance of advertising-level profits. At the same time, the copyright maximalists didn't *quite* win this round. We'll almost certainly get somewhat better access to several million OP books—and will have to hope (and work to see) that the price (monetary and otherwise) isn't too high.

---

## Masthead

*Cites & Insights: Crawford at Large*, Volume 9, Number 4, Whole Issue 114, ISSN 1534-0937, a journal of libraries, policy, technology and media, is written and produced by Walt Crawford, Director and Managing Editor of the PALINET Leadership Network.



*Cites & Insights* is sponsored by YBP Library Services, <http://www.ybp.com>.

Opinions herein may not represent those of PALINET or YBP Library Services.

Comments should be sent to [waltercrawford@gmail.com](mailto:waltercrawford@gmail.com). *Cites & Insights: Crawford at Large* is copyright © 2009 by Walt Crawford: Some rights reserved.

All original material in this work is licensed under the Creative Commons Attribution-NonCommercial License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/1.0> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

URL: [citesandinsights.info/civ9i4.pdf](http://citesandinsights.info/civ9i4.pdf)