S

ometimes it feels as if the Google book digitizing project has been with us all of our professional lives, and the Google Book Settlement not much less than that. In fact, Google began its digitization program early in the decade; the Authors’ Guild lawsuit was launched in 2005, and the Settlement was proposed in late 2008. A great deal of the Settlement story remains to be written, to say nothing of its impact on many individuals, organizations, society, and the future of books. It seems that everyone connected to learning and knowledge production—not just attorneys and authors—has a view about Google’s activities, offerings, and strategies. The Google Settlement has even become a topic of conversation among strangers seated next to one another on long airplane flights.

Is the Google digitization program a boon for all, leveling the playing field for have-not users, fostering creation of new knowledge, accelerating the research process, saving libraries money, and much more? Is it a fiendishly crafty way for Google to make even more money than so far? Recently Michael Cairns, former president of R.R. Bowker and Managing Partner, Information Media Partners published a paper titled “Database of Riches: Measuring the Options for Google’s Book Settlement Roll Out.” In it he estimated that Google’s annual subscription revenue for licensing to libraries could approach $260 million by year three. Some of our airplane conversations address matters such as:

• Will Google enjoy a monopoly unlike any heretofore seen in the publishing world?
• Will Google exploit its tremendous digital assets within the blessing of legal precedent?
• Will other businesses rush to do the same, following Google’s precedent, thus introducing useful competition into the digital marketplace?
• Will Google enhance the role of libraries as the new destination? Will libraries become passé?

Many have views about these matters, and Against the Grain thought to solicit and represent some of them for its readers. Here you will find Ivy Anderson’s opening piece, in which she argues the benefits of having millions of books available to readers and responds to several key librarians’ and scholars’ concerns (such as long-term preservation). Pamela Samuelson offers an author’s viewpoint, based in her legal expertise; as a classical scholar, James O’Donnell values Google, but he does not want to lose the added (and enormous) metadata value that librarians add. The international view is not often heard in the United States, so ATG invited Paul Whitney to write from the perspective of one of the U.S.’s major partners, Canada, our neighbor to the North. Stuart Hamilton provides a broad international view, grounded in his work as policy advisor for IFLA. Finally, Jonathan Band has given us permission to reproduce his “March Madness” flow chart, showing possible paths forward for the Settlement. Perhaps by the time this ATG issue is published, some of the paths forward will be clearer, but by no means will the Judge’s ruling be the end of the story. You will read more in these pages as time passes. And you will be better informed for your next plane trip!

---

Ann Okerson (Associate University Librarian, Collections & International Programs, Yale University)
<ann.okerson@yale.edu> • http://www.library.yale.edu/okerson

---

Against the Grain / June 2010

Ann Okerson

---

BORN AND LIVED: Hallein, Austria; lived in US, Canada, and UK.

IN MY SPARE TIME I LIKE TO: Pursue, wherever I may be, the following: dark chocolate, cupcakes, French macaroons (I bought my first computer in the 80s to create a database of chocolate shops).

FAVORITE BOOKS: For spare-time reading, I enjoy mysteries, particularly with an international setting. I could list many, but here are a few favorites: Colin Cotterill (Coroner Dr. Paiboun, Laos); James Church (Inspector O, North Korea); Henning Mankell (Sweden or Africa); Eliot Pattison (Inspector Shan, Tibet); Lisa See (China); Xiaolong Qiu (Inspector Chen, China). Well, let me stop here and hope at least one of these authors is new to you. And email me for more suggestions!

PET PEEVES: Drivers speaking on cell phones. Young people smoking. Don’t they know the statistics? People who need to speak on cell phones wherever they may be and don’t care who listens to whatever they’re saying (sometimes really private stuff).

MOST MEMORABLE CAREER ACHIEVEMENT: Starting NERL, the NorthEast Research Libraries Consortium.

GOAL I HOPE TO ACHIEVE FIVE YEARS FROM NOW: Less obsessive about working so much of the time? But, then, I do love it!!

HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS: For the librarians, I see enormous opportunities to deliver information to our readers in many new, different ways and to adopt a flexible, innovative mindset. Things we do will keep changing—no standstills. I think (hope) we will be visibly closer to solutions for our digital preservation, so that the shift to e-content needn’t be so worrisome for scholars and librarians. This will allow us to think about how best to manage and consolidate our physical collections. Many traditional physical library spaces will be repurposed or else shrink/go away. The library will be wherever readers are. We will be five years closer to the vision of a universal digital library, as each of our institutions continues to play its part in making that wonderful future happen.

---

<http://www.against-the-grain.com>
Hurtling Toward the Finish Line: Should the Google Book Settlement Be Approved?

by Ivy Anderson (Director of Collections, California Digital Library, University of California, Oakland, CA 94607)

Digitization Partnerships: The Promise and the Peril

Like many of the objectors, participating libraries went through their own period of outrage and indignation when details of the Settlement first came to light. What? We would have to buy back access to our own books? Why did Google let us down in abandoning its fair use defense? Why should the parties be allowed to create an artificial revenue model for works that are long out of print, books that would no longer exist at all outside of used bookstores, if the libraries themselves hadn’t purchased and maintained them at great expense over decades and indeed generations? How can they do this without our agreement as to terms, since it is we who have made these books available to them in the first place? Hasn’t our stewardship paid for these books many times over? Isn’t this why copyright law contains unique exceptions for libraries, in recognition of our mission to further the public good? Wasn’t the appropriate use of our own copies in light of fair use principles our decision to make?

The problem with this view, of course, is that libraries did not initiate this enterprise, and we are not its only beneficiaries. The Google project placed two sets of commercial interests as loggerheads, with copyright law in the middle. Admittedly, libraries took a risk in engaging in a partnership so legally entangled. But let’s be honest: though few seem willing to admit it, revitalizing the world’s heritage of books for a digital age — a task that many considered impossible only a few short years ago — appears within reach today almost entirely due to Google’s enterprise vision. Even the Open Content Alliance, which CDL joined a year before becoming a Google partner, was in some sense a response to GBS (although it had other important antecedents, as well, thanks to Brewster Kahle’s equally inspired vision). When Google’s competitors withdrew their support for that project, no other funders stepped in to fill the breach. The plain fact is that despite the idealistic adjurations of some, the resources required to digitize our cultural book heritage on a grand scale are not likely to be marshaled in the U.S. by libraries and the public sector alone.

At CDL, we’ve done some estimating of what it would take to convert the roughly 15 million unique books in University of California library collections to digital form absent the Google enterprise, using the best alternative technology available today. The answer? Half a billion dollars, and one and a half centuries. And that is just the University of California’s books.

I like to compare this to the building of the great Temple of the Sagrada Familia in Barcelona, a city with which my family has an ancestral connection. When my husband’s grandmother left Barcelona as a young girl in the late 19th century, the Sagrada Familia had barely erected its first stone. In 2006, more than 125 years later, her great-granddaughter traveled to Barcelona for the first time, where she was able to observe Gaudi’s monumental edifice, still under construction. At this writing, completion is projected for 2026.

The speed at which Google is converting this content is not without costs of its own. Google’s iterative approach to building large-scale services has drawn criticism from some scholars accustomed to work that is honed and polished before it is released. This is, in part, an argument about means, not ends. Like those progressive JPEG images that start out blurry on the screen and become sharper as the details fill in, Google’s services are improving over time as it continually upgrades and enhances its images and metadata. Over time we will be able to replace those missing or still-blurry pages with better versions. Where the value of the content warrants it, we can selectively invest in more meticulous rendering, textual markup, and other enhancements.

Two cases are illustrative here. CDL has digitized a large number of public domain books with the Internet Archive, some of which have also been digitized in our Google partnership. Although CDL had to suspend its Internet Archive book scanning project earlier this year after Microsoft withdrew its support and additional grant funding proved elusive, we have every expectation that we will take up comparable projects with Internet Archive in future, because its technology is better suited to certain types of uses (better artifactual rendering, for example). The Early English Books Online (EEBO) database

(The image presented here is a mashup: for the actual book in HathiTrust, see http://hdl.handle.net/2027/uc1.b543888)

Recently, Google and the plaintiffs filed their final briefs in defense of the Google Book Amended Settlement Agreement (ASA) that is before the New York Southern Federal District Court. As the rhetoric around the Settlement heats up to white-hot intensity, I’d like to offer a few personal thoughts from my vantage point at the California Digital Library.

The University of California Context

The University of California Libraries are Google’s second-largest library digitization partner; we are also the second-largest book digitization partner of the Internet Archive, thanks to generous funding in the past from Microsoft, Yahoo, the Alfred P. Sloan, and Kahle/Austin Foundations, and other sponsors. In all, UC Libraries have now digitized 2.5M books from their collections through these projects, both in- and out-of-copyright. Within the University of California, some of our closest faculty colleagues are also among the Settlement’s most prominent critics.

In our view, the proposed settlement is hard-ly perfect; as Google acknowledges in its brief, it’s a compromise among parties with differing agendas and motivations. CDL is a staunch supporter of the underlying aims of the Google Book project to make the knowledge enshrined in the world’s great libraries discoverable and accessible across the globe, and we support the public benefits that will ensue, including the benefits to libraries, if the Settlement is approved. At the same time, public criticism has been good for the Settlement, producing very real improvements in the amended version that is now before the court — improvements that would not have been made without that criticism. Long live democracy!
Hurtling Toward the Finish Line
from page 18

marketed by ProQuest is another example in which, through an innovative partnership with libraries and scholars, basic scans are enhanced with detailed markup for a subset of carefully-selected works.

In the meantime, a great deal of value is already being derived from the Google work as it stands today. Students and scholars report finding much formerly-hidden material, journalists and etymologists are mining its content for historical information, and even some of Google’s severest critics have said that they can no longer imagine life without GBS. This is neither an either-or proposition nor a zero-sum game. All of these services are fulfilling a niche, along with libraries, in a new information ecology that we are only now beginning to understand even as we participate in its unfolding.

Settlement Pro and Con

So in the long run, is the Google Settlement a good thing, or a bad thing? Before answering that question, let’s look at just a few of the major criticisms that have been levied against the Settlement.

The Institutional Subscription will become too expensive because it has no meaningful competition. Well, it’s hard to know that, of course. In fact, we don’t even know today how many books it will contain, nor what the scan or OCR quality of the content will be, given the variability of the overall corpus. But we do know that there are at least three checks on the institutional subscription price that should mitigate price-gouging. First, the broad distribution requirement in the Settlement’s dual objectives means that prices cannot become so high that few choose to subscribe. Second, libraries themselves are savvy evaluators and negotiators of online content who can be expected to evaluate this offering rigorously and skepticaliy, and to eschew a subscription unless the price is acceptable for the benefit derived. Since none of us knows how our users will engage with this material, these assessments ought to be conservative. Third, the provisions for pricing arbitration built into the agreements between Google and the participating libraries will allow them to challenge price increases that they deem unwarranted; a provision that is intended to be exercised not on the basis of narrow self-interest among a small set of contributing libraries but on behalf of all libraries.

Academic authors want to release their books, not see them locked up. Indeed, no disagreement here; and the amended Settlement now explicitly provides for this (according to Google and the plaintiffs, this was always possible, but in the ASA it is now called out). We intend to work proactively with rights holders who would like to enable broader access to their books and to develop mechanisms that can help to make this straightforward.

The Settlement will give Google a monopoly over orphan works and is anti-competitive. It’s hard for me to see how Google’s activities to date can be viewed as anti-competitive when GBS is almost single-handedly responsible for the eBook explosion that is swirling all around us, with new entrants popping up every day. That may be a controversial assertion, but eBooks and eBook readers were a languishing backwater until Google stimulated the market by putting books online through its library and publisher partner programs. If anything, Google’s entrance into the retail space is likely to engender fiercer competition. It seems cynical at best for rival behemoths Microsoft and Amazon to decry Google’s impending monopoly over a sliver of the eBook market — much of it of uncertain commercial value — under the noble-sounding rubric of the Open Book Alliance. But then, competition makes strange bedfellows. As to orphan works, the Settlement should, if anything, goad us all the more toward a legislative solution. It is as risksome to me, as it is to other critics, that Google should be uniquely empowered to collect royalties on behalf of absent rights holders who may have long ago relinquished any economic interest in their works. Still, the ASA addresses this in a far more satisfying manner than its predecessor. Finding a better long-term solution to the orphan works problem is something we all can get behind.

When the purposes that we first envisioned when embarking on these projects — all arguably fair uses of this content — are reviewed against the Settlement impacts, it’s hard to view the Settlement as anything but a positive development. More books will be available in full view, both to libraries and to consumers. New services will be developed for print-disabled users and for large-scale computational analysis, further unlocking digitization’s transformative potential. Disclosure of rights information through a central registry (at least for U.S. books) is likely to have far-reaching impacts, facilitating the eventual orderly release of books into the public domain. Google’s competitors are likely to join the push for orphan works legislation, increasing its chances of success. And with the Settlement behind us, we can all proceed in an environment of greater certainty.

What if the Settlement is not approved?

For libraries, failure of the agreement would hardly be a crisis. The benefits that we initially envisioned — improved discovery and full-text search of our vast legacy collections, and broad public availability of works that are out-of-copyright or otherwise released by their copyright owners — will still be realized. The fears of some Settlement objects — of monopolistic pricing and the forced commercialization of materials that are long out-of-print — will melt away like the elusive Vancouver snow. Participating libraries may still choose to undertake novel services, without the unwelcome restrictions imposed by the Settlement. As long as Google and others continue to partner with us, we will go forward in reinvigorating our collections for a new digital age.

The Google Settlement is fundamentally about whether Google and rights holders will be allowed to implement a particular set of business models for a certain set of books. I believe the Settlement should be approved, because it will create new and valuable services for libraries as well as consumers. But many of Google’s participating libraries have their own plans for these books, plans that do not ultimately depend on the outcome of the Settlement. The greatest risk for libraries if the Settlement is not approved is that further legal setbacks might lead Google to abandon its interest in library digitization altogether. If that were to happen, a unique opportunity would be lost that is not likely to be repeated in our lifetime.

Life Beyond Google Book Search

What of our relationship to the Google Book project itself? Some of the concerns we hear from faculty have nothing to do with the Settlement per se, but rather with the long-term implications of GBS for library collections and services. Let me close with a few words about some of those concerns.

To our scholars who worry that we are about to throw our physical collections overboard in favor of digital surrogates of sometimes uneven quality, I want to say: not to worry. True, libraries everywhere find themselves having to consign more and more of their physical collections to remote storage as campus space grows increasingly scarce and user preferences migrate online. And some libraries — the UC’s far less than others — are addressing the space crunch by de-accessioning low-use materials that are widely held with the knowledge that they can borrow these items from another library if need be. (Many cooperative initiatives are now underway to share such information and ensure that enough copies are retained throughout the nation’s system of libraries to protect the integrity of the scholarly record.) That train has already left the station, and it’s happening independently of large scale digitization. What digitization offers is a valuable complementary mitigation strategy: we can now make those remote collections eminently browsable, saving time and expense both for users and for libraries. As a library user, you can now determine whether that book is really what you’re looking for before you request it, not afterward — and in some cases, the digital surrogate may indeed be all that you need. Libraries can promote these “hidden” volumes more effectively to their users, while limiting delivery costs to just those items that are truly wanted. This browsable and/or searchable digital surrogate — which is the quality level that most of the Google mass digitized scans are aimed at — is not a replacement for the original print book, and was never intended to be.

To our scholars who worry that we are outsourcing our library collections and services to Google, again I want to say: please don’t worry on this score, either. For one thing, our mission as stewards of the cultural record, we who have opened up our collections to digitization are shoudering this role with vigor. While Google and others are making these books discoverable online to a general audience, the University of California, along with other
peer institutions, is creating a robust shared access and preservation service for our mass digitized books, one that adheres to professional standards, through our partnership in a groundbreaking enterprise called the HathiTrust. If you haven’t heard of HathiTrust yet, you soon will. No UC library user need go to Google to search the full text of our books, or to find accurate bibliographic information, or to view and download those that are in the public domain: she can go to http://catalog.hathitrust.org/ and be reassured that those books will be there, in ever-improved versions, for the long-term. HathiTrust now numbers 5.4 million volumes from 26 libraries and is growing at a rapid rate, all searchable, all viewable if in the public domain (or otherwise rights-cleared), and all designed to inure to the long-term benefit of the nation’s libraries and their users. The digital library of the future resides not with Google, but with us. And we are building it today.

At the same time, Google, Internet Archive, and others, are providing an invaluable service in bringing the vast holdings of the great research libraries to a worldwide audience and integrating that content with general-purpose internet search services and other content. As one colleague has written, “Who among us has not benefited from a Google search?” In participating in these efforts, we are fulfilling our long-standing public service mission. The Google Settlement, if approved, will further these aims by providing more content, in more ways, to an even wider audience.

But in the end, approval of the Settlement is not a make or break event for libraries. Despite the claim that the Google Settlement promises to build “the greatest library in history,” libraries are not leaving the future of information to Google and these other partners alone. Nor need we wait, Godot-like, for fugitive national legislation to begin the work of serving up our cultural heritage in digital form. Through a combination of efforts, including public-private partnerships such as that of libraries with Google, we can go forward in this transformative enterprise together.

This piece was initially published in Anderson’s blog, at: http://www.cdlib.org/cdlinfo/2010/02/16/hurtling-toward-the-finish-line-should-the-google-books-settlement-be-approved/.
An Academic Author’s Perspective on the Google Book Settlement

by Pamela Samuelson (Professor, Berkeley Law School & School of Information, University of California, Berkeley, CA 94720-4600)

During the long-awaited “fairness hearing” about the proposed settlement of the Authors Guild v. Google lawsuit on February 18, 2010, I was one of the 21 non-party objectors or opponents of the proposed settlement to whom the judge granted five minutes to present their views. After introducing myself and noting that I had filed two letters objecting to specific terms of the GBS Settlement, the latest one on behalf of 150 academic authors, I made the following points.

Most of the books that will be regulated by the settlement agreement are out-of-print books from the collections of major research libraries such as the University of California, and most of these books were written by scholars for scholarly audiences.

Many scholars own copyright interest in their books and inserts at least for electronic distribution. Many of them have clauses in their contracts that allow author reversion rights upon the book going out of print. Most of these books will be core parts of the institutional subscription database that will be licensed to universities such as my own.

In the past year I have spoken to many colleagues at U.C. Berkeley and elsewhere about the proposed Settlement. When I asked them whether they would be willing to allow their out-of-print books to be made available on an open-access basis, each has said yes. Academic authors tend to believe that orphan books should be available on an open access basis too.

Orphan books are not a trivial problem. The Financial Times has estimated the number of books likely to be orphans as between 2.8 to 5 million. These books will form a core part of the institutional subscription database to which my university and others are expecting to subscribe.

The Plaintiff’s memorandum responding to objections characterizes open access advocacy as “a prime example of...parochial self-interests” (p.3). The memo goes on to say that the interests of academic authors conflict with those of the author subclass. It also bears mentioning that academic authors would not have brought this lawsuit against Google, because we tend to think that scanning books to make snippets available is fair use. If this case goes back into litigation, instead of being settled, I will be writing briefs in support of Google, not in support of the Authors Guild.

But it’s not just I and the 150 people who signed the supplemental academic objection letter who endorse open access for these books. Last August, a letter was sent to the judge on behalf of the UC Academic Senate — representing 16,000 faculty members at the University of California — which expressed concern that the open access preferences of academic authors would not be respected.

More important, though, is the open access recommendation of the U.S. Copyright Office in its report on orphan works. The Copyright Office considered and rejected an escrow model for orphan works akin to that in the amended Settlement agreement. Once the orphan status of a work has been determined, the Copyright Office recommends that the work should be available for free use. Congress modeled its orphan works legislation on the Copyright Office’s recommendation. With all due respect, I believe that what should be done about orphan works is a public policy issue that should be decided by Congress, not private parties or the courts.

It is far more consistent with the utilitarian principles of copyright law to allow orphan books to be made freely available once we know that they are, in fact, orphaned. This is important to academic authors, because the Plaintiffs only care about maximizing revenues for the millions of orphan books that will be in the institutional subscription database. Academic authors want to maximize access. This is why it is crucial to include meaningful constraints on price hikes as part of the settlement agreement.

There is a fundamental difference in perspective between the Plaintiffs and academic authors about why books are valuable. For the Plaintiffs, books are commodities to be exploited for maximum revenues. However, for academics, books are more like a slow form of social dialogue. Books from the past open the conversation that scholars pick up and carry forward. The books academics write earlier in the conversation and set the stage for the conversation to be carried on by our successors.

The set of objections I made on behalf of academic authors should not be swatted down one by one, as they were in the Plaintiff’s Objection memo; instead, they should be viewed as important component parts of the cultural ecology of knowledge in academic communities. This ecosystem will be impaired if the ecosystem envisioned in the settlement agreement is adopted, instead of the one that has long prevailed and should prevail in the future for academic communities.

While I could live with the GBS Settlement if it was amended as I have suggested, I worry very much about the precedent that would be set by approval of this particular Settlement.

Google’s founders say that the company’s goal is to organize all of the world’s information. As we all know, books are not the only information resource that contains the world’s information. I have been wondering for some time which sector of the copyright industry will be next to have its works scanned by Google for inclusion in its search database.

If this settlement agreement is approved, Google and possibly others may feel free to go out and scan other copyrighted works. And if their rights holders object, the pragmatic response might well be: hey, we could litigate about this, but I have a good fair use defense, and it would be expensive and ugly to litigate, so why don’t we just reach a deal on my terms right now? Approval of the settlement would give Google unfair leverage in such negotiations.

But beyond that, I believe that approval of this settlement would encourage other class action lawsuits, which would then seek to justify their efforts to remake copyright law by saying, in effect, “Congress is too dysfunctional to address this problem, so we must be allowed to do it through a class action settlement.”

For Pamela Samuelson’s home page, where all her writings on the Google Book Settlement are posted, see: ischool.berkeley.edu/~pam. The full transcript of the fairness hearing is available at: http://thepublicindex.org/. Also well worth consulting is the lively synopsis of the hearing by James Grimmelmann at: http://laboratorium.net/archive/2010/02/20/ gbs_fairness_hearing_report.

Pamela Samuelson is recognized as a pioneer in digital copyright law, intellectual property, cyberlaw and information policy. She has written and spoken extensively about the challenges that new information technologies present for public policy and traditional legal regimes. Since 1996, she has held a joint appointment with the Berkeley Law School and the School of Information. She is the director of the Berkeley Center for Law & Technology, serves on the board of directors of the Electronic Frontier Foundation and the Electronic Privacy Information Center, and on advisory boards for the Public Knowledge and the Berkeley Center for New Media. She is also an advisor for the Samuelson Law, Technology and Public Policy Clinic. Since 2002, she has also been an honorary professor at the University of Amsterdam.
There are three questions we must ask about a project for mass digitization:

First, should the work be done?
Second, should the results be made widely available?
Third, can the results be easily used?

I propose to answer these questions in order. The first I will answer gratefully, the second cheerfully, and the third grumpily. At a moment when everyone has an opinion about the Google Book Settlement, I may not seem grumpy enough to many, but my gratitude and my cheerfulness are real.

First, yes, the work should be done. Ten years ago, I was privileged to lead a National Academy of Sciences study group looking at the digital future of the Library of Congress (our results were published in book form by the National Academies Press, LC21: A Digital Strategy for the Library of Congress [2000], a volume that I think still holds up quite well). That extraordinarily talented team of scientists, scholars, and librarians was positively wistful about a possibility we imagined beyond our reach. “Congress should just appropriate a billion dollars and digitize the whole thing,” one of our number liked to say, and we all shook our heads: this couldn’t possibly happen, not in the real world of Congress and politics.

Well, a private company has gone and done it. Not all of it, not a billion dollars worth, but the quarter of a billion dollars or so that Google is said to have spent on this task is an extraordinary benefaction to the public weal. We should work hard to remember to say thank you. If they hadn’t done it, nobody else was fighting to the fore to do it first, least of all Congress.

The second question is a trick, of course. Should the works be made widely available? Of course. But once that is said, the knives will come out. It should all be made widely and freely available. Information still wants to be free, access should be open, and rainwater should be beer. That is not quite happening, though it is worth noticing that as far as Google is concerned, the material over which they have complete say, the U.S. out-of-copyright material undoubtedly before 1923, is in fact being made freely available. The negotiated settlement with the publishers took longer than anyone imagined, is complicated beyond the capacity of non-specialists to remember from one week to another, and does indeed put a price tag on the in-copyright material in the project.

This is not Google’s fault. The intellectual property laws we have and the publishing industry to which they have given rise are real choices made in a free society and, indeed, we have all benefitted mightily from them. That there may be better ways to make laws and do business is obviously a topic for discussion and debate, and the wrangling proceeds. But for now, the Google Settlement as proposed reminds one of Churchill’s crack about Democracy as the worst form of government — except all the others that have been tried. The Google Settlement will open vast ranges of published material to a broad public for easy use. It’s a start. (And of particular note is that the Settlement will make it possible to have access to so-called “orphan works” — in copyright but owned by individuals unknown and not safely reprintable. This too is progress.)

So I am cheerful again, and not a little grateful. Impatient, as well: if the legal wrangling will end, I’ll be able to get my hands on things now tantalizingly just beyond my reach.

Third question: can the material be easily used? This is another trick, of sorts. Making the material available free of charge might be nice, but there’s much more to accessibility than price tag. Information has to be findable and, when found, usable. The reader needs what one of my colleagues on that National Academy study called “intellectual access”. In a mass of information, the reader has to be able to find the right material at the right time and read it and think about it and even copy chunks of it.

Here’s where Google’s most notable failing strikes the eye. Numerous observers, perhaps most flamboyantly the linguist Geoff Nunberg, have called attention to the woeful defects of Google’s metadata. Google trusts that search engines will gain all the access one needs and has accordingly shown astonishingly cavalier neglect towards one of the great achievements of human intellect: I mean that body of highly structured and awesomely accurate metadata embodied in the catalog records of our libraries. Yes, those metadata have their limits. Yes, searches can find you much that the old MARC records cannot find. But those records remain astonishingly useful for many purposes. Each MARC record reduces a volume or cognate set of volumes to a consistent, structured, and easily intelligible representation in very small compass. When there are masses of information to be dealt with, those representations are far easier to use than a collection of search hits from the original material can ever be. But Google has thrown those data aside.

How does this affect a real working scholar? I will give just one example. If I want to find the seven volumes of Tyrrell and Purser’s famous edition of Cicero’s letters, published in the first decades of the last century, I have a very particular goal in mind. I want to find Volume One, then Volume Two, then Volume Three, and so forth, until I have found Volume Seven. If I were in a physical library, I would expect to find them together on a shelf, or else in a record (keyed to the MARC record) of their present temporary location. In Google Books, I think I find several copies of Volumes One and Two, and no more. Furthermore, I find nothing to help me know whether I am looking at first editions or revised editions, and nothing to tell me where to look for Volumes Three through Seven. Those books are still very useful to me: but not in Google.

Now Tyrrell and Purser cross the boundary into possibly-copyright, but there are not (when last I looked) any indications that Volumes Three through Seven have shown up on Google’s scope at all. But I may simply be wrong.

This story repeats itself throughout the corpus of scholarly literature that I am curious about. I am squirming away on my laptop huge PDF files of Google books versions of things I want to read — because I’m never sure I’ll find them again, and because I’m frustrated that I can’t find complete sets of things that were published even on the same day together.

One could make everything in Google Books absolutely free to me today and it wouldn’t help this problem. The digital representations that Google has made available aren’t yet a library — and indeed, in an important sense, they aren’t even books yet, any more than a stack of de-accessioned volumes in a dumpster are books any longer. Books are books when they are alive and speaking, their contents known and knowable. Books do furnish a room, but care must be taken with the room and with the books. Google has a long way to go on that score.

And that’s why, and where, and how I’m grumpy.
The Google Book Settlement: Canadian Perspectives

by Paul Whitney (City Librarian, 350 West Georgia Street, Vancouver, British Columbia, Canada V6B 6B1)

A

Canadian analysis of the Google Book Settlement (GBS) must be placed in the context of a cultural policy, which has always taken a protectionist stance largely motivated by the perceived danger of cultural incursions from south of the 49th parallel. While several years of a minority Conservative government have signalled a move to a more free market approach to the regulations governing cultural industries, sensitivities are still present and quickly manifested if threats to cultural sovereignty appear. Amazon’s announced intention to open a Canadian based warehouse to service Amazon.ca was front page news in our national newspaper in February 2010, with industry representatives denouncing the threat this would pose to Canadian booksellers. The virtual existence of Amazon.ca supplying its products (Canadian and foreign) through a subsidiary of Canada Post was approved by the government several years ago, on the basis that there were no employees in Canada and, therefore, it fell outside the regulations designed to protect Canadian distribution and retail. This counter-intuitive perception that the creation of Canadian infrastructure and employment by a foreign owned company already fully serving Canadian authors was a threat to Canadian interests is indicative of the confusing outcomes which can arise from “brick and mortar” regulatory regimes applied to virtual enterprises. The ongoing debate over Amazon.ca indicates that Canadian cultural protectionism is alive and well.

Given this environment, it was especially surprising that Canadians were largely absent from the initial debate on the Google Book Settlement (GBS). Unlike governments in France and Germany, the Canadian government took no position on the GBS before the U.S. Court or in the media. English Canadian publishers largely signed on to the Book Rights Registry and took no formal position prior to the September 2009 Court deadline for submissions. Their motivating factor appeared to be that if there was money to be made, they needed to be part of the initiative. In keeping with their European counterparts and in contrast to their Anglophone counterparts, ANEL (Association Nationale des Éditeurs de Livres), urged their members not to sign on to the Registry but appear not to have made a submission to the Court. As far as I can ascertain, there were only three Canadian submissions by the September deadline:

• A strongly worded objection from the Canadian Standards Association, as a publisher owning international copyright rights including U.S. publications. CSA described the GBS as “anticompetitive, arguably violates antitrust laws, and improperly uses the class action mechanism...to force a perpetual business deal upon class members for the future use of copyrighted works in ways that go well beyond the facts that gave rise to this lawsuit in the first place.” The brief concluded with the CSA’s concern for the future implications of the GBS: “Google will likely continue its practice of ‘copy first, settle later’ and, after its monopoly power is firmly entrenched through this action, will likely attempt to leverage an even better deal for itself at copyright holders’ expense next time.”

• The Canadian Urban Libraries Council (CULC) offered “general support in principle” for the GBS, arguing that without it “the probability of a subscription-based service with this vast body of work being available outside the United States is very unlikely.”

• The Writers’ Union of Canada (WUC) Statement of Objections indicated support for the GBS establishment of the Book Rights Registry, while expressing a number of concerns over “expropriating the copyrights of foreign rights holders” if their works were published only outside the U.S. without authorization for U.S. distribution. Other concerns raised by the WUC included:

— The settlement should not permit future digitization by Google without voluntary sign up with the Book Rights Registry.

— Google should not be permitted to license and profit from orphan works in the absence of U.S. Congress legislation on the matter.

— Libraries and non-profit higher educational institutions should be required to pay a licensing fee to provide public access to the database, and digital copies should not be used by them to replace titles that are commercially available. The WUC Chair was quoted as saying on the free access to the database through public libraries, “That really sticks in our craw because we think it could have copyright implications in Canada.”

— Authors of foreign works should be represented on the Books Rights Registry.

— Rights holders should have the choice of opting into new uses not covered by the GBS.

Canadian stakeholder engagement with the GBS increased significantly following the filing of the Amended Google Book Settlement (AGBS) in November 2009. The AGBS limited the scope of access to digitized works (but not their actual digitization) to works published in the U.S., UK, Australia, and Canada, countries described as having common legal heritage and similar book industry practices. Paul Atken, Executive Director of the Authors Guild, one of the main plaintiffs in the case against Google, stated that this narrowing of coverage in the AGBS meant “Ninety-five percent of foreign language works are out,” meaning that “the lion’s share of the potential unclaimed works are now out of the Settlement.” What does make Canada unique in this grouping is its active French language publishing sector. While reliable Canadian publishing statistics are elusive, it is reasonable to assume that 25% to 33% of Canadian publishing is French language. It is interesting to speculate if the anomaly of full database access to these “foreign” language titles might generate additional revenue relative to European-published French language titles.

The two major national English language Canadian publisher associations, The Association of Canadian Publishers (membership comprises 133 Canadian owned and controlled publishers) and the Canadian Publishers’ Council (18 publishers including foreign-owned trade and education publishers and legal publishers), both issued general letters of support for the AGBS shortly after its release. The Canadian Publishers Council specifically noted its satisfaction with the change to the definition of “commercially available,” which reads in the AGBS (new text underlined):

‘Commercially Available’ means, with respect to a Book, that the Rightsholder of such Book, or such Rightsholder’s designated agent, is at the time in question, offering the Book (other than as derived from a Library Scan) for sale new from sellers anywhere in the world, through one or more then-customary channels of trade to purchasers within the United States, Canada, the United Kingdom, or Australia.

Publisher pleasure with this amendment is understandable, as the distinction between “in print” and “out of print” is significant in the AGBS, with Google restricted for in print titles from displaying more than text snippets and publishers controlling the right to sell full text. In this age of Internet book selling, the definition change means effectively that any book available anywhere in the world is deemed to be in print worldwide and Google has restricted rights on what it can do with the book.

As Canadian publishers quickly fell into line in support of the AGBS, the mobilization of writers and educators in opposition to the deal started to coalesce.

An online petition from writers opposed to the AGBS circulated, and by early January 2010 had 250 signatories. The WUC submission to the U.S. Court in 2009 was described by a spokesperson for dissident writers as failing to
take an official position and trying “to work it from the inside.” The writers under the name “Canadian Writers Against Google Settlement” filed an objection to the AGBS to the U.S. Court on January 28th asking that Canadian copyright holders be removed from the agreement. Several new arguments (from Canadians at least) against the settlement were introduced:

- As well as violating the Berne Convention (an argument made forcibly by European interveners), the agreement would be in violation of U.S. obligations under NAFTA
- Canadian authors' moral rights would be violated under the agreement
- Competition and privacy concerns should be addressed
- Canadian provisions for addressing orphan works should be respected
- Canada's bi-lingual and bi-juridicial heritage and tradition set it apart from the other countries included in the AGBS

As was the case with the VUC, the Union des Ecrivaines et des Ecrivains Quebecois, the primary Quebec writers organization, did not advise members on a specific position on the AGBS.

The Canadian Association of University Teachers (CAUT), representing over 65,000 teachers, librarians, and other academic staff, also intervened with the U.S. Court on the AGBS in late January. CAUT echoed a number of the objections raised by other Canadian groups, including that the AGBS is in conflict with international copyright and trade agreements, ignores Canadian legislation on moral rights and orphan works, is in conflict with the separate Quebec legal and commercial regulatory regimes, and includes minimal privacy protections. CAUT also introduced the objection that the interests of its members are at odds with those of the AGBS plaintiffs in that “academic authors generally place a higher premium on access than is reflected in the (AGBS).”

As we await the next stage of the ongoing GBS saga, from a Canadian perspective it is difficult to imagine that it could be implemented as written without it leading to transformative change in Canada’s regulatory, publishing, and library environments. Whether the transformation is catastrophic or liberating or a little of both remains to be seen and will certainly be in the eyes of the beholder. As a librarian I tend to “fetishize” access (in the memorable phrase of European critic Roland Reuss) and am inclined to agree with CULC in its assertion that implementation of the GBS is a necessary first step in providing universal access to our print heritage, while providing reasonable protections for writers and content providers. I worry that “universal access” for a number of years will be limited to the United States, and that there has not been enough consideration of the research imbalance this will create, especially if institutional subscriptions are constrained in any number of ways for institutions outside the U.S. Setting aside the implications for academic research, the image of a Canadian having to travel to a U.S. public library to access a digital text of a Canadian title is both troubling and offensive. The impression left in a June 2009 meeting between Google representatives and Canadian educators and librarians that GBS implementation was at least ten years away in Canada does not offer much hope in this regard.

The only thing that is certain is that this process will not get any easier as it proceeds. I do believe, however, that the imperatives of the emerging digital reality will make a resolution to the multifaceted tensions surrounding the GBS both necessary and desirable for all concerned. An outcome that only addresses English language content must be seen as a partial and interim solution.

---

**Endnotes**

2. Ibid., p 5.
3. Letter from Jeff Barber, CULC Chair, to Judge Chin of the U.S. District Court for the Southern District of New York, August 31, 2009.
E ver since Google began digitizing millions of books in 2002, the Google Book project has fascinated the international library community. The tantalizing possibility of universal access to a massive number of books from American and European libraries, with further expansion to institutions elsewhere in the world — this is the stuff of librarians’ dreams. Even as the years have gone by, and more books have been digitized, at the same time louder voices are heard against the Google initiative. The idea of universal access seems to have faded somewhat from librarians’ minds, even if the possibilities Google Book offers remain attractive and seemingly within reach.

The International Federation of Library Associations and Institutions (IFLA) is the leading international body representing the interests of library and information services and their users. Founded in 1927, IFLA has 1600 member associations and institutions in approximately 150 countries around the world. In its 83-year history, IFLA has authored and published many books, and therefore has a great interest in the resolution of the Google Book question. Furthermore, some IFLA members are partners in the digitization programme itself, and as such are keen to see the success of the project and increase access to their collections.

As is well-known, in 2005 Google was sued by the Author’s Guild and the Association of American Publishers for illicit digitization of copyrighted works. In October of 2008, the three parties negotiated an out-of-court settlement that has since been amended and is now awaiting approval by a court in New York. Both author and publisher of books, IFLA falls within both sub-classes of plaintiffs in the suit, and consequently has formalised its position on the settlement in an official statement prepared by its Copyright and other Legal Matters (CLM) Committee and approved by its 20-person international Governing Board.1 This statement was submitted as an amicus brief to the court in September 2009.

IFLA’s position is that of a deep commitment to the principles of freedom of access to information and the belief that universal and equitable access to information is vital for the social, educational, cultural, democratic, and economic well-being of people, communities, and organizations. In light of this, we welcome Google’s potential contribution to achieving these goals by providing access to a digital library of millions of books. The proposed settlement under discussion in New York could prove very fruitful. However, we also believe that there are some clear consumer and access issues that must be satisfactorily addressed before we can support this, or any similar subsequent agreements among libraries, rightsholders, and corporate partners.

Chief amongst these is the issue of territoriality. As an organisation representing libraries all over the world, IFLA has strong concerns about the territorial limits of the settlement. As it stands, the expanded services permitted under the settlement would be provided only to users located in the United States. Users outside of the USA will only have access to a more limited version of the Book Search service. If the Google Settlement is approved in the United States and if Google is not willing or able to reach agreements with rightsholders in other countries, the consequence will be an ever-widening inequality in access to books in digital format. IFLA wants library users worldwide to have the widest possible access to information via the Internet.

Further to this, we are also concerned about the monopolistic nature of the project. During the IFLA World Library and Information Congress in Milan in 2009, IFLA’s CLM Committee held a session on the settlement that included a presentation by a Google representative. This presentation revealed that Google has digitized 10 million books (and is proposing to digitize an additional 20 million) at a cost of c. $750 million. The immensity of the project, and the fact that Google has a five-year lead, makes it challenging for others to start viable competing projects. In consequence, a large proportion of the world’s heritage of books in digital format could be under the control of a single corporate entity, should the settlement be approved.

Monopolistic concerns also contribute to our thoughts on the pricing policy proposed in the Settlement. The economic terms for the Institutional Subscriptions Database will be governed by two objectives: (1) the realisation of revenue at market rates; and (2) the realisation of broad access by the public, including institutions of higher education. IFLA members’ recent experience has been that publishers of scientific journals have prioritised revenue generation over broad access, forcing many libraries to cancel subscriptions. If the beneficial societal effects of Google Books are to be fully realised, it is critical that the importance of broad access be given strong weight in the Settlement.

Libraries will pay an as-yet undisclosed fee to license access to the database. In view of the potential monopolistic nature of the project, and the collaborative manner in which it must be implemented, IFLA believes that libraries must have an integral — and not merely advisory — role both in the establishment of pricing for the Institutional Subscriptions Database and the manner in which revenue from it is allocated to the parties, including libraries.

It is unclear if libraries as consumers can negotiate on behalf of their users, and they apparently cannot negotiate access through consortial arrangements. It must therefore be possible for any library or institutional subscriber to request the court to review the pricing of services provided.

In connection to this, IFLA would like to see an emphasis on the role of libraries as providers of content, as well as users or consumers. Librarians must be involved in the policy setting process for the Book Rights Registry, because libraries serve as the contributors of content to the database, and as the primary consumers of content on behalf of their users. Libraries’ massive investments in collecting, organizing, and preserving this corpus are as essential for the project’s success as the work of the authors and publishers who created the stock in the first place.

Connected to pricing policy is an area we have a great deal of concern about, and something that libraries all over the world are contending with on a regular basis when offering access to digital resources. In copyright, contracts too often override statutory exceptions and limitations in ways that diminish users’ rights. The Settlement should, therefore, clearly state that nothing in it supersedes legislated users’ rights, including specific and general exceptions for libraries and their users, and any existing or new approaches to making orphan works accessible.

IFLA’s amicus brief also highlighted the possible censorship issues in the proposed Settlement. Google may exclude from the database 15% of scanned books that are under copyright, but out-of-print. This could exclude one million books. Google is likely to come under pressure from interest groups and even governments to exclude books that are purported to contain “undesirable” information. If Google submits, this could lead to the suppression of these books worldwide and the stifling of freedom of expression. IFLA therefore believes it is of the utmost importance that the settlement obliges Google to publish lists of books that are excluded from its services, and the reason for the exclusion.

Finally, patron privacy is such a core value for libraries that a court order is usually required to force a library to disclose individuals’ use of library resources. Some of the services to be offered under the proposed Settlement imply that Google will collect and retain information about users’ activities. However, the Settlement does not specify how users’
privacy will be protected. IFLA has urged the U.S. court to require Google to cooperate with library associations and other representatives of users’ interests to ensure that adequate measures are taken to protect personally identifiable information.

Across the pond, the European Union has been considering the implications of the Settlement, and European library organisations such as the European Bureau of Library, Information and Documentation Associations (EBDLIA), and the Association of European Research Libraries (LIBER) have produced their own position statements. On the September 7, 2009, IFLA, EBLIDA, and LIBER, along with other library representatives, appeared at a special hearing at the European Commission in Brussels to comment on the potential effects the settlement would have for Europe and the rest of the world. Like the other plaintiffs in the Settlement, more than six months later we are still waiting to discover the decision of Judge Denny Chin. What happens next will not only be crucial for citizens of the U.S., but also for students, scholars, and library users in the rest of the world, as the first possible steps towards access to a global digital library are either taken or held back pending further amendments.

Effective with January 2008, Stuart Hamilton was appointed as IFLA’s first Senior Policy Advisor. His Ph.D research examined freedom of access to information on the Internet worldwide, and the ways in which libraries can overcome barriers such as censorship or the digital divide to ensure that library users receive the best possible access to online information resources. Hamilton has lectured extensively around the world on these and related matters, and his findings have been widely published. Prior to accepting this position, Hamilton worked for IFLA’s FAIFE Office (2001-2006).
Now that the fairness hearing on the Google Book Settlement has occurred, it is up to Judge Chin to decide whether the proposed settlement is “fair, reasonable, and adequate.” This chart attempts to diagram some of the possible paths forward.

Notwithstanding the complexity of the chart, it does not reflect all the possible permutations. For example, it does not mention stays pending appeals nor whether litigation would proceed as a class action. Moreover, the chart does not address the substantive reasons why a certain outcome may occur, e.g., the basis for Judge Chin accepting or rejecting the Settlement. And it doesn’t begin to address the issue of Congressional intervention through legislation. In short, the precise way forward is more difficult to predict than the NCAA tournament. And although the next step in the GBS saga may occur this March, many more NCAA tournaments will come and go before the buzzer sounds on this dispute.