# From the University Presses — Google 2.0: Still a Mixed Blessing?

Column Editor: **Sanford G. Thatcher** (Director, Penn State Press, USB 1, Suite C, 820 N. University Drive, University Park, PA 16802-1003; Phone: 814-865-1327; Fax: 814-863-1408) <sgt3@psu.edu> *www.psupress.org*

In the wake of a tidal wave of objections filed to the original **Google Settlement** in the last few months leading up to the "fairness hearing" that was planned for October 7, the parties to the agreement prevailed on **Judge Denny Chin** to allow them time to revise it and submit a new version by November 9. They missed that deadline but made an extended deadline, presenting **Google 2.0** to the court in literally the 11th hour of Friday the 13th.

Objections had come from many quarters, ranging from private citizens to companies like **Amazon.com** to foreign governments, but among the most compelling were those presented on behalf of the academic community, by the **U.S. Justice Department**, and by **Register of Copyrights Mary Beth Peters**.

**UC-Berkeley** law professor **Pamela Samuelson** was a leading voice among academics, writing of the "audacity" of the Settlement for her **Huffington Post** blog on August 10 and challenging it on both anti-trust and representational grounds (*http://www.huffingtonpost.com/pamela-samuelson/the-audacity-of-the-googl_b_255490.html*). Much of her argument was repeated at greater length in a very articulate and persuasive letter dated August 13 and signed jointly by 21 faculty leaders from the **University of California**, who "constitute the entire membership of the Academic Council, the executive body of the Academic Senate, and the chair of the Academic Senate's Committee on libraries and Scholarly Communication" (*http://bits.blogs.nytimes.com/2009/08/17/uc-professors-seek-changes-to-google-books-deal/*). They grouped their concerns under three main headings: "Risks of Price Gouging and Unduly Restrictive Terms"; "Support for Open Access Preferences"; and "Privacy and Academic Freedom Issues." The letter makes a particularly compelling statement about how the Settlement takes no account of the interests that academic authors have that are different from those of members of the **Authors Guild**, which took upon itself the role of representing the entire class of authors. "Specifically, we are concerned that the **Authors Guild** negotiators likely prioritized maximizing profits over maximizing public access to knowledge, while academic authors would have reversed those priorities. We note that the scholarly books written by academic authors constitute a much more substantial part of the Book Search corpus than the **Authors Guild** members' books." I think the same point could be made by university presses about how well the **Association of American Publishers** represented their interests in negotiating the Settlement. Our priorities, too, are different from those of **McGraw-Hill**, **Pearson**, et al.

The Justice Department, while recognizing the significant public benefit that the Settlement could bring from its "potential to breathe life into millions of works that are now effectively off limits to the public," also took the Settlement to task for its inadequacy of class representation, but focused attention on the disadvantaged positions of foreign rightsholders and authors of out-of-print books (*http://searchengineland.com/department-of-justice-files-objections-to-google-book-search-settlement-26144*). The Settlement's provisions allowing **Google** to negotiate with the **Book Rights Registry (BRR)** for new derivative uses of out-of-print titles and paying unclaimed funds to rightsholders who had opted in to the Settlement prompted this objection in the Department's brief: "There are serious reasons to doubt that class representatives who are fully protected from future uncertainties created by a settlement agreement and who will benefit in the future from the works of others can adequately represent the interests of those who are not fully protected, and whose rights may be compromised as a result." The Department also raised two main questions about anti-trust implications of the Settlement: "First, through collective action, the Proposed Settlement Agreement appears to give book publishers the power to restrict price competition. Second, as a result of the Proposed Settlement, other digital distributors may be effectively precluded from competing with **Google** in the sale of digital library products and other derivative products to come."

Finally, in a hearing before the House Judiciary Committee on September 10, **Mary Beth Peters** characterized the Settlement as "not really a settlement at all, in as much as settlements resolve acts that have happened in the past and were at issue in the underlying infringement suits. Instead, the so-called settlement would create mechanisms by which **Google** could continue to scan with impunity, well into the future, and … create yet additional commercial products without the prior consent of rights holders. For example, the settlement allows **Google** to reproduce, display and distribute the books of copyright owners without prior consent, provided **Google** and the plaintiffs deem the works to be 'out-of-print' through a definition negotiated by them for purposes of the settlement documents. Although **Google** is a commercial entity, … the settlement absolves **Google** of the need to search for the rights holders or obtain their prior consent and provides a complete release from liability. In contrast to the scanning and snippets originally at issue, none of these new acts could be reasonably alleged to be fair use." Because the settlement, in effect, "is tantamount to creating a private compulsory license through the judiciary," it is "the view of the **Copyright Office** [that] the settlement proposed by the parties would encroach on the responsibility for copyright policy that traditionally has been the domain of **Congress** [and] we are greatly concerned by the parties' end run around legislative process and prerogatives…. Moreover, the settlement would inappropriately interfere with the on-going efforts of **Congress** to enact orphan works legislation in a manner that takes into account the concerns of all stakeholders as well as the United States' international obligations." (For a link to the full testimony, see *http://laboratorium.net/archive/2009/09/10/gbs_marybeth_peters_written_testimony*.) The Settlement, in short, serves as an insurance policy for **Google** to pursue its project of digitizing what **Dan Clancy**, Engineering Director for **Google Book Search**, has estimated to be "between 80 and 100 million books in the world" free of any liability for the vast majority of those books, which are out of print. No other commercial competitor of **Google** would have such sweeping legal protection to conduct its business, which a compulsory license approved by **Congress** would create for all.

The **Amended Settlement Agreement (ASA)** takes significant steps in responding to many, though not all, of the objections raised. For academic authors who are rightsholders and opt in to the Settlement, it provides the opportunity to set prices at zero or to use **Creative Commons** licenses for designating kinds of uses that require no payment or permission. While the Settlement, in restricting its geographical scope to include only works registered in the U.S. or published in Australia, Canada, and the United Kingdom, provides for representation on the **BRR** board of an author and publisher from each of these three foreign countries, there is no guarantee that any academic author or publisher will hold such a seat.

*continued on page 57*

**Biz of Acq**
*from page 55*

everything they need to perform at their best. It is important that the planners do all that is in their power to make this expectation a reality. Planners need to give themselves some time to relax and quench any qualms they might have about the programming. If this is done, all of the hard work and effort will lead to a fulfilling and successful program that all attendees will enjoy and appreciate!

**From the University Presses**
*from page 56*

The Justice Department's concerns about representation were also met by the **ASA** with a partial response. Works published in non-English speaking countries were taken off the table with a stroke of the pen. That change solved one major problem but created others. It significantly limits the potential value of the **Google Book Search** database by excluding millions of works published outside the four countries. (One wonders about the omission of New Zealand, which is the home to several universities presses at the universities of Auckland, Canterbury, Victoria, and New Zealand. Is this another indication of the lack of academic representation in the Settlement?) It also exposes **Google** potentially to suits by authors and publishers in other countries for the original scanning of their books from the participating libraries. **Google's** argument about "fair use" has no obvious basis in the laws of these foreign countries as it does under U.S. copyright law, and it could be legally challenging for **Google** to prevail in their courts.

With regard to authors of out-of-print books, the **ASA** tweaks the definition of what is "commercially available" in a variety of ways and, most significantly, creates a new "**Unclaimed Works Fiduciary**" (**UWF**) to assume some of the responsibilities in representing these authors' interests that were originally assigned to the **BRR**. But, as **Randal Picker** of the **University of Chicago Law School** points out in his perceptive working paper titled "Assessing Competition Issues in the **Amended Google Book Search Settlement**" posted on November 16 (*http://ssrn.com/abstract=1507172*), the **UWF** only offers a partial solution: "The **UWF** mechanism enables separate representation of those interests. But the settling parties have limited the role of the **UWF** to merely stepping into the shoes of the registry in some circumstances. They could have broadened the role for the **UWF** to have the **UWF** step into the shoes of the rightsholders of unclaimed books instead. Had that been the focus, the **UWF** would then be an elegant solution to the going forward problem of how to license the orphan works." In addition to providing this new mode of representation, the **ASA** also specifies different uses of unclaimed funds: none will go to other rightsholders, but instead a portion can be used after five years to help the **BRR** cover the cost of locating rightsholders and what is left over after ten years may be distributed by the **BRR** to charities focused on improving literacy with the approval of the court and in consultation with participating libraries.

The **ASA** deals with anti-trust issues in a number of ways. Perhaps most crucially, it accepts that the court's approval of the Settlement will not result in automatic immunity for it from anti-trust challenges in the future. It thus postpones resolution of whether or not pricing provisions will prove to have anti-competitive effects in the marketplace. That change responds to fears about monopolistic power in part. Another change in this direction is the excision of the much criticized "most favored nation" clause that would have guaranteed **Google** the same terms as offered to any possible competitor by the **BRR**. Still other changes speak to fears of price-fixing and foster more flexibility: the pricing algorithm used to set default prices for the consumer purchase of books will be controlled by **Google** alone, not as previously in conjunction with the **BRR** and rightsholders; **Google** may discount book prices at its discretion and will allow other companies like **Amazon.com** to sell access to the **Book Search** titles for consumer purchase as well; **Google** and rightsholders may negotiate a different split of revenues for any title included in any of the authorized programs from the 37/63 designated in the original Settlement. On the other hand, the virtual monopoly that the Settlement provides to **Google** as a sole-source provider for out-of-print books remains unchanged despite the addition of the **UWF** to the **BRR** as a potential licensor to third parties "to the extent permitted by law." As **Randal Picker** observes: "My understanding is that **Google** does not believe that that provision actually enables either the registry or the **UWF** to license the works to third parties and that they instead believe that legislation would be required by **Congress** to make that operative. Be very clear: the settlement agreement is giving **Google** rights directly to use the orphan works. **Google** is not getting rights to the extent permitted by other law."

So, how does this **ASA** meet the needs of academic libraries and university presses? Libraries, which are not direct parties to the Settle-

*continued on page 58*

ment but are obviously affected in many serious ways by it, have not gained much with the changes made. Little was done in response to concerns about privacy beyond a clause that holds **Google** responsible for not sharing personally identifiable information about users with the **BRR** "absent valid legal process." The **BRR** board will still not have anyone on it representing libraries. Concerns about price-gouging were not directly addressed and, to the extent they are connected with price-fixing worries, were partially alleviated but mainly delayed to resolution at some future time. The scope of content has been further restricted by the redefinition of "book," which now excludes any work that contains more than 20% of pages containing musical notation as well as books reproduced in microform, comic books, calendars, and compilations of periodicals. Illustrated books are even further disadvantaged by the **ASA's** excluding children's book illustrations from the definition of "insert." Maybe the most significant step forward for libraries (but only public not academic libraries) is that the **BRR** is now free to authorize additional public-access terminals beyond the one per building originally provided.

For university presses, undoubtedly the most salient change is the introduction of the possibility of negotiating the 37/63 revenue split for every title in every program. Concomitant with this benefit, of course, is the additional transaction costs that any publisher will incur in negotiating with **Google** title by title. Also at play here is what **Google** will decide to do with its publisher partners whose contracts will now all need to have an addendum that will permit **Google** to move ahead with its **Google Editions** program, which will run parallel to the **Book Search** program and be based on agreements signed with the current partners that choose to expand their dealings with **Google** in this way. **Ian Paul** of *PC World*, in announcing this new initiative in a blog on October 16 (and congratulating **Google** for eschewing DRM so that the books supplied through its program can be read on many different types of devices, not just dedicated e-book readers), noted: "Under **Google's** payment scheme, publishers will receive about 63 percent of the gross sales, and **Google** will keep the remaining 37 percent. **Google** also hopes to offer **Editions** titles through other online book retailers. In this scenario, online retailers would get 55 percent of revenues minus a small fee paid to **Google**, and publishers would get 45 percent. **Google** may also create deals to sell **Google Editions** books directly through a book publisher's Website, but no details have been announced for how that scenario would work" (*http://www.pcworld.com/article/173789/google_editions_embraces_universal_ebook_format.html*). Now that the 37/63 split has been made open for negotiation in the **ASA**, one wonders how publishers will react to a non-negotiable split for **Google Editions**. The split is peculiar to begin with as it relates to no traditional standard in book publishing. Whence did it come? According to **Ken Auletta** in his new book about **Google**, this split is traceable to **Google's** experience with linked Website advertising, where **Google's** share includes a 15% administrative charge and makes it total about 37%. This transposition of a revenue-sharing model from one domain to an entirely different one is questionable at best and seems purely arbitrary. To impose it unilaterally on publishers of all kinds seems equally arbitrary. So, too, the 45/55 split between publishers and online retailers. That mimics what is standard for trade-book publishing, but hardly represents well the main business that university presses conduct, where "short" discounts of 20% to 25% for monographs and textbooks are more common. It will be interesting to see how presses evaluate the pros and cons of working with **Google** through either the Settlement arrangement or the partner program; perhaps some will experiment with both. The Settlement imposes significant extra costs on presses in burdening them with requirements to negotiate with individual authors over such matters as display percentages, and it also deducts an administrative fee that might be as much as 20% or 25% from the 63% of the gross that is due publishers, leaving them in effect with not much greater a share of the overall income as **Google** gets. Much will depend, therefore, on whether **Google** decides to be flexible at all with the 37/63 split in the partner program. Uncertainty also exists for the pricing algorithm that the **ASA** now mandates **Google** alone to define and control. According to the Memorandum of Law submitted by the parties to the court on November 13, "the Pricing Algorithm will be designed to simulate how a Rightsholder would unilaterally price its Book in a competitive market." (Links to this Memorandum and other **ASA** documents may be found here: *http://lawprofessors.typepad.com/law_librarian_blog/2009/11/amended-google-book-settlement-filed.html*.) Well, how do rightsholders determine what price to charge? In the eBook world, in fact, this is not an easy question to answer. Many publishers are struggling with it now, and some are even inclined to set the price for a book differently depending on what kind of platform offers it for sale and what the features of that platform are. (For a very suggestive discussion of this question in relation to the introduction of the **Kindle 2 Reader**, see the comments by **Tony Sanfilippo**, Sales and Marketing Director at **Penn State University Press**, on the Press's blog here: *http://psupress.blogspot.com/2009/04/kindle-2.html*.) If this kind of variable pricing by device and feature becomes prevalent, it will pose a huge challenge to **Google** in making its pricing algorithm truly reflective of what is happening in the marketplace.

**Google 2.0** is unquestionably an improvement on **Google 1.0** in many respects, and the chances for approval by the judge after the final fairness hearing now scheduled for February 18 now seem much better than before. But, besides the loose ends and only partly satisfactory solutions identified above, the Settlement still leaves much to be desired in other respects. Although it is good to have some funding explicitly aimed at helping identify and locate the rightsholders of unclaimed, including orphan, works through the redirection on monies not claimed by rightsholders, publishers in general and university presses in particular continue to face the daunting challenge of knowing what rights they actually have. As **Mike Shatzkin** observed in his blog about "A serious issue for big publishers" on April 14, "they are largely in the dark about what rights they own…. The **Google**-related issues primarily revolve around whether the rights to an inactive book (or, in the settlement lingo, what they would call 'not commercially available') have reverted to the author or are still held by the publisher. Publishers also have problems with books on which they unambiguously have the rights to print and sell copies. What they don't know, without looking at the original contract, is whether the language in it gives them a shot at an ebook, a print-on-demand edition, or allows them to include some of the material in that book in an electronic database. Even looking at the contract might not tell them if they have the rights to use artwork that is in the book in any other edition" (*http://www.idealog.com/blog/a-serious-issue-for-big-publishers*). Some commercial publishers face an additional challenge that university presses fortunately do not have to worry about: companies that once were independent have merged, sometimes several times over, and tracking the disposition of rights across various stages of merger can be a major obstacle to clarity about who now holds what rights. But university presses have the same problems commercial publishers do with rights reversion and old contracts not containing any or inadequate language about electronic rights.

The **BRR** plays a central role in the whole Settlement scheme, yet it is faced with an enormous challenge of creating a sophisticated technical infrastructure to record rights claims and process payments to **Google**, rightsholders, and potential third-party licensees. As one who has witnessed the growth of the **Copyright Clearance Center** as a member of its broad of directors for nearly twenty years, I have a special appreciation for what is required to be successful in this kind of business. It requires an organization nimble on its feet, always seeking new ways to serve its customers better, and a large and dedicated staff who have the public interest at heart. The **CCC** is now over thirty years old, but the **BRR** is expected to get up to speed almost overnight by comparison. Related to this is the sorry state of the metadata that publishers have so far had to work with in getting ready to claim books in the **Google** database. One can only hope that the **BRR** will be able to make marked improvements in the metadata once it is off and running with a full staff. Otherwise, publishers will continue to be burdened with yet another type of heavy transaction cost in just getting their books properly set up in the system.

Finally, there is the continuing concern about content, not only that the **Book Search** database will ill serve the needs of people who want to access illustrated works such as art history books but also that the quality of the content it can deliver is not high. Numerous

*continued on page 59*

**From the University Presses**
*from page 58*

critiques have displayed the results of the often erratic nature of the scanning that **Google** contractors have performed, complete with smudges, misaligned pages, and even pages containing images of the scanners' thumbs. But the problems go beyond simple quality of reproduction. There is a serious concern about metadata here, too, from a scholar's point of view. As **Geoffrey Nunberg** so devastatingly catalogued in his article for *The Chronicle of Higher Education* (August 31) titled "**Google's Book Search**: A Disaster for Scholars," the current metadata "are a train wreck: a mishmash wrapped in a muddle wrapped in a mess." **Nunberg's** survey covers errors in dates, problems with classification, and mismatches of titles and texts. I particularly sympathize with his critique of **Google's** decision to use BISAC codes to classify books. "Why," he wonders, would **Google** "want to use those headings in the first place"? As **Nunberg** notes, "The BISAC scheme is well-suited for a chain bookstore or a small public library, where consumers or patrons browse for books on the shelves. But it's of little use when you're flying blind in a library with several million titles, including scholarly works, foreign works, and vast quantities of books from earlier periods. For example, the BISAC Juvenile Nonfiction subject heading has almost 300 subheadings, like New Baby, Skateboarding, and Deer, Moose, and Caribou. By contrast, the Poetry subject heading has just 20 subheadings. That means that **Bambi** and **Bullwinkle** get a full shelf to themselves, while **Leopardi**, **Schiller**, and **Verlaine** have to scrunch together in the single heading reserved for Poetry/Continental European. In short, **Google** has taken a group of the world's great research collections and returned them in the form of a suburban-mall bookstore." For most university press books, I can attest, the BISAC codes compel one to be very creative in trying to use enough codes to represent the subject of a scholarly book at all adequately. Just to give one example, there is no way of straightforwardly identifying a book about modern Latin American politics. One has to cobble together a set of codes covering History/Latin America/General, History/Modern/20th Century, and Political Science/Government/Comparative at a minimum. And to identify a book in feminist philosophy, one has to leave the category of philosophy altogether to find any code representing feminist or gender studies (under the main rubric of Social Science). **Google's** decision to employ BISAC codes is yet one more glaring revelation of how skewed the Settlement is toward the interests of trade-book authors and commercial trade-book publishers rather than academic authors and academic presses. And the irony of it all is that the vast majority of books now among the ten million **Google** has in its database are academic books, making **Book Search** a potential boon for scholars everywhere — if only **Google** had talked with the right publishers to begin with!

# Group Therapy — A Case of Discredited Research

Column Editor: **Jack G. Montgomery** (Associate Professor, Coordinator, Collection Services, Western Kentucky University, Bowling Green, KY) <jack.montgomery@wku.edu>

***Column Editor's Note:*** *I posted this question to* ***COLLDV-L*** *and received a host of thoughtful answers that span the range of opinion on this complex issue. I sincerely thank all those who weighed in on this question. A similar issue has arisen concerning* ***Disney's Baby Einstein*** *product. (See the* ***New York Times*** *10/23/09 issue* ***http://www.nytimes.com/2009/10/24/education/24baby.html?_r=1)*** *however, the following answers concern the* ***Belles*** *book. — **JM***

**GRIPE:** Submitted Anonymously. In the September 2009 issue of ***Against the Grain*** was an article by **Steve McKinzie** of **Catawba College** entitled "The case for getting rid of a celebrated book." It his article, **McKinzie** discussed the discredited title *Arming America: The Origins of a National Gun Culture* by **Michael Belles** which was first given the **Bancroft Literary Prize** in 2001. Later in 2002, the prize was withdrawn and the author discredited due to professional scholarly misconduct with regard to the research and its presentation. **McKenzie** made the case for removing such a book from the library's collection. Although I understand **McKenzie's** argument, I am personally confused as to what our responsibility is in such matters. I would like to hear from other librarians but would like to remain anonymous. Can you help me?

**RESPONSE:**
Submitted by **Linwood DeLong** (Collections Coordinator, **University of Winnipeg Library**, Winnipeg, MB, Canada)

I am a Canadian and therefore possibly not totally qualified to weigh in on this one, but because it is an intriguing topic, I will do my best.

To me, the issue should be first and foremost, the quality of the books in our collection. If we discovered that a history book about any topic was full of factual errors, based on faulty research, citing phantom sources, etc. then we would remove the book for those reasons. We remove many old books because they contain outdated information — a book about the U.S. that refers the "48 states and their capitals" would disappear from our shelves, unless it were a famous travel book, such as **De Tocqueville's** accounts of his travels.

Books that take a controversial stand - we had a recent, highly publicized case in Canada about a book published by **McGill Queen's University Press** that took a very controversial stand about native peoples' issues — are different. Our library, probably many libraries, bought the book, because it presented this viewpoint and would enable students to study the articulation of the viewpoint and respond to it. At the far end of this spectrum are completely nonsensical books (we all see promotions for self-published books) that are so un-scholarly that they are not useful at all in our collections. We don't buy those.

We probably have some books in our collection that deny that the Armenian genocide ever occurred. Many of us would dispute this, but propaganda material (if it is clearly understood to be so) can still be useful, again for study and research purposes.

I'm starting to stray a bit from the topic. If we had *Arming America* in our collection, or a book about a medical topic in which the results were demonstrated to be false because of the use of phantom data or the deliberate misuse of existing data, I would argue for the removal of the book from our collection.

I guess that I am trying to draw a line between factual inaccuracies, misrepresentation of data, etc. and controversial opinions. It appears, from what I saw in the e-mail on **COLLDV-L**, that *Arming America* is of the first type.

I enjoy collections development problems or challenges and would be pleased to respond to others, if you think that my response is useful.

**RESPONSE:**
Submitted by **Sarah Tusa** (Associate Professor, Coordinator of Collection Development & Acquisitions, Mary & John Gray Library, **Lamar University**, Beaumont, TX)

First of all, I must admit that I am not familiar with the details of the complaints against the author's research conduct or methodology, but it would seem that the validity of the information presented in the book was very probably tainted by the improper research and invalid presentation of the research results, then that book is very similar to an outdated edition of any other book. If the author were to produce a revised (and corrected) edition, we would definitely withdraw the original edition. Some larger, more comprehensive (probably **ARL**) libraries might make the argument to keep the original, tainted edition as a part of publishing history. However, I personally would be tempted to withdraw the *Arming America* book even without the prospect of getting a new, revised edition, for the same reason that we withdraw out-of-date medical books: We at least ***attempt*** to minimize the amount of outdated or invalidate and/or discredited information that our students can get their hands on in our library.

I don't consider it censorship, if the errors in research and in the end product are documented. As an integral part of a teaching institution, the library has a responsibility to

*continued on page 60*