From the University Presses -- The Google Settlement: Boon, Boondoggle, or Mixed Blessing?

Sanford G. Thatcher

Penn State Press, sgt3@psu.edu

Follow this and additional works at: https://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation
Thatcher, Sanford G. (2009) "From the University Presses -- The Google Settlement: Boon, Boondoggle, or Mixed Blessing?," Against the Grain: Vol. 21: Iss. 1, Article 36.
DOI: https://doi.org/10.7771/2380-176X.2519

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
From the University Presses — The Google Settlement: Boon, Boondoggle, or 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

Everyone seems to agree that the Google settlement announced in October 2008 represents a milestone of some kind in the development of access to information, but there is a wide spectrum of views about whether, overall, this is a good thing or a bad thing as far as the general public interest is concerned. Publishers appear to be as mixed in their opinions as librarians.

A lively debate is ongoing over the liblicense listserv on the merits of the settlement. Rick Anderson, in a posting on January 23, prefers to accept the positive: “Look at what the Google settlement has done: the general public now has far better (though still imperfect) access to vastly more literary and scientific writing than it ever has had before. This access is, by any sane definition of the term, free. (More comprehensive access is available at a price, but what’s available at no charge is still amazing.) Even better, the content to which we now have access is, for the first time ever, fully searchable, and we can get it from our homes and around the clock. Better still, the public has paid virtually nothing in return for what it now gets.” To the skeptics, he says: “Sometimes I think we’ve actually made an art out of letting the perfect be the enemy of the good.” Ann Okerson, in her posting on December 17, also finds “commendable” the Google Book Search pilot project. The goal of this project is “to explore upstream metadata capture and enhancement using publisher and vendor ONIX metadata”. Centralized federations of metadata are but one of many potential solutions to improving metadata. Another is the Book Industry Study Group (http://www.bisg.org/) and their ONIX Data Certification Project (http://www.bisg.org/documents/certification_projectdata.html).

NISO’s goal is to build understanding among the variety of players in this process of transforming metadata to fulfill the needs of the many different users and uses in the chain. The subtleties of differences in needs and the significant infrastructure investments made by different constituencies make it unlikely that the community can settle on one single data structure or transport mechanism. What is potentially more likely is creating standardized crosswalks and application profiles for different standards used in the community. Obviously, standards or best practices will play a role in the eventual solutions or improvements to the exchange of metadata. However, just as important will be a deeper understanding of the investments and the strengths that each participant in the exchange process brings to the table. Each constituency will have something to learn from the others in the chain, which might help reduce costs and improve functionality for everyone.

Against the Grain / February 2009
From the University Presses from page 69

Paris, which choked off free trade in books [and, not coincidentally, spurred the movement to adopt copyright legislation as an antidote to monopoly power]. Google is not a guild, and it did not set out to create a monopoly…. But the class action character of the settlement makes Google irreplaceable to copyright owners. As Chris Castle, a former assistant to Hipster writing from the UK in The Register in a posting titled “Monopoly Money from Digital Books” (http://www.theregister.co.uk/2008/12/31/chris_castle_google_books_and_beyond/), elaborates: “If a competitor tried building a competing book registry by negotiating licenses for in-copyright works, they would have to bear the startup costs — and the cost of licensing. If the competitor is rewarded for respecting authors’ rights by obtaining favorable terms, that advantage can be taken away by Google. Why? Because one of Google’s goodies from its dominant position in the settlement negotiation is ‘most favored nations’ price protection. The registry is contractually required to offer Google any better terms it if it gives anything using any data or resources that Google provides the registry, or that is of the type that Google provides. So even if a competitor wants to build a parallel infrastructure from scratch, and wasn’t using any of Google’s data or any resources that they would have been paying for through Google’s MFN. There is no advantage in ‘doing it right’ except a clear conscience — an MFN inhibits competition.” Castle warns ominously that this monopoly might well not stop at books, quoting Google co-founder Sergey Brin as seeing the new book registry as the first step toward monetizing “other areas of digital media, like video.” As Richard Johnson notes in The Christian Science Monitor (December 23, 2008), “the proposed deal not only solidifies Google’s dominant position in Internet search, it gives the franchise a virtual monopoly on the long-tailed out-of-print book market.” And even though public-domain works would be offered free to the public, the mere fact that access to them will be restricted under the settlement to Google searching alone means that “in effect, for the one-time price of a scan, Google now proposes to secure and enforce a monopoly on the digital texts of works that belong to the public” — a situation that he clearly considers deplorable. As he succinctly summarizes the situation, the “settlement is a stark reminder that markets are sustained by many different motivations than libraries. Control over library collections, once guided by the values of learning and research, is now a commercial matter. Goodbye free, hello fee.”

Instead of settling with authors and publishers, what if Google had pursued its suit over fair use to its legal conclusion in the courts, as many in academe had hoped when Google initially position itself as a champion defender of the principle? Some noted copyright authorities, like Larry Lessig and William Patry (the authors of the leading text on fair use, who is now employed by Google), believe that Google would have prevailed on the merits of the argument. Others, like Siva Vaidhyanathan, had their doubts. So do I. When one considers the Ninth Circuit whose authority on copyright issues stands as one of the most important in the country, and the author of the Fair Use Standard in the Harvard Law Review (March 1990) that identified “transformative use” as the “heart” of copyright law in a sense quite different from how the Ninth Circuit has interpreted it, I think the odds were against Google prevailing. Instead of continuing to fight what would have been a losing battle, and having ended up being the worst in a very close battle with at least an uncertain outcome at best, Google struck a deal for a modest investment of $125 million that is likely to be paid back many, many times over in future revenues. According to Lessig, “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 than it would have had.” Vaidhyanathan’s verdict is that “fair use in the digital world is just as murky and unpredictable (not to mention unfair and useless) as it was yesterday.”

Whatever the implications for fair use may be, the question remains whether, realistically, there was any alternative to relying on the private sector to accomplish transformative digitization. Darnton, among others, thinks “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. By spreading the cost in various ways — we could have provided authors and publishers with a legitimate income, while maintaining an open access repository or one in which access was based on reasonable fees. We could have created a National Digital Library…. Others are not so sure. James O’Donnell, posting on lblogcon on January 23, provocatively noted that had something like the state of play around LC on these issues a decade ago, and the prospects for public funding in support of such a project were slim, to say the least.” Some public funding has gone into efforts like the Million Book Project, which received grants from the National Science Foundation totaling $3.6 million to help with digitizing children’s library books, but it has taken eight years to reach this level, and is yet very far from becoming the Universal Library it had the ambitions to become. It has been brought under the umbrella of the Internet Archive, itself established in 1996 with similar ambitions, but none of these other projects, or all of them together, have come close to reaching the level of digitization that Google has achieved in a much shorter period of time. As Paul Courant observed on his blog, “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 create a new and vibrant economy where in some cases book by book, and publisher by publisher. I’m confident that we would have gotten there in time, serving the interests of all parties. But in time ‘would surely have been many years.” Others credit Google with having given a tremendous boost to efforts within academe that can build on what is now possible. As scholarly communication librarian John Wilkin writes in Library Journal (December 23, 2008) that the cooperative HathiTrust Project, launched initially within the CIC libraries but now involving California’s, Virginia’s, and other universities’ libraries, too, with the “aim to create nothing short of a universal digital library,” has found the Library Google Project to be “integral in seeding HathiTrust with a large body of materials as well as inspiring a new level of digitization activity by libraries, library consortia, and other partners, such as the Open Content Alliance.” And he notes also that “libraries are tight, the threat of action for some collaborative activities that the original Google agreements with libraries did not and that are at the core of what HathiTrust wants to accomplish.

The very name HathiTrust connotes that a private enterprise like Google ultimately cannot be trusted with the mission “to protect the historical record and to ensure its future for the public.” “Google,” Wilkin says, “cannot be that trust for the public.” But one wonders whether the will even exists in universities to make such a commitment, perhaps with help from state and federal governments and from foundations, necessary to achieve control just of the intellectual property immediately produced by their own faculty and to make it freely available to the public in the way advocates of infrastructure for universities proclaim others, had they chosen to do so, to capture this developing market for themselves. But this “missed opportunity” was allowed to pass, and librarians have spent decades now ruing the consequences. In principle, there seems to be no reason that a “grand alliance” of the nonprofit kind Darnton limns couldn’t yet be formed to meet the challenge Google monopolized, although he thinks “it is too late now.” But even if the means exist, as they once did for publishing STM journals en masse in a non-profit manner, the will does not seem to be there to make a challenge to the Google monopoly possible. Universities appear to be content to rely on the market even when their rhetoric suggests otherwise. If they weren’t, wouldn’t those some one hundred U.S. universities that support presses be more willing to allow them to make all their publications “open access” instead of continuing to require the presses to recover 90% or more of their costs from sales in the marketplace? When push comes to shove, and budgets are tight, the public interest in the public interest always seems to defer to deference on market mechanisms to make the system of scholarly communication work.

So, forced as they are to rely on the market to cover most of their operating costs, how do university presses view the Google settlement? I think it is fair to say that opinions among press directors vary as much as opinions among librarians do. While presses generally were excited about the new possibilities for selling backlist titles opened up by Google Book Search and its facilitation of “the
worthy that, except for illustrations in children’s books, the settlement excludes images from the scope of the agreement altogether — and hence only “mostly” solves the problem.) Depending on how one evaluates the potential monetary value of out-of-print books, attitudes toward the possible financial benefits from the various programs that the settlement envisions for Google to launch, beginning with institutional subscriptions and extending through sharing of ad revenues and supplying print-on-demand editions, range from the optimistic to the skeptical. I wonder myself how much demand there will be for this vast sea of out-of-print material. There is, after all, good reason these books went out of print in the first place: demand simply had deteriorated to the point where offset printing technology made reprinting uneconomical. Books with strong continuing value have never gone out of print, whether classics of philosophy like Hume’s Treatise on Human Nature, foundational works in social science like Morgenstern and Von Neumann’s Theory of Games and Economic Behavior, popular expositions of science like Einstein’s The Meaning of Relativity, or great novels like Austen’s Pride and Prejudice. As a publisher for forty years of scholarly works in the humanities and social sciences, I have read many works that have now outlived their usefulness, either because their theoretical frameworks have long since been superceded or because their factual information has been corrected by later investigation; many of them are of interest now only to people who are writing about the history of disciplines, and even these investigations would likely focus on the primary works that had achieved near classic status in these fields (the “paradigm-changing” works, to use Kuhnian language), rather than the multitude of “case studies” in the social sciences or applications of various popular interpretative approaches like deconstruction in the humanities. Some old books really do deserve to be left in the dustbin of history. Thus I count myself among the skeptics about how great the financial returns will be from this monetization of the out-of-print corpus. Still, I have been pleasantly surprised at how well the “long tail” has worked so far for older backlist titles — though not yet producing much more income than eBooks have for most publishers, namely, less than 5% of total revenues — and I am prepared to be pleasantly surprised again at the eventual results the Google settlement might produce.

On the other hand, with all the benefits, actual and potential, come some significant costs. What Google will charge for its services — 37% of all revenues generated under the programs envisioned under the settlement — seems excessive. It is nearly double, for example, what most literary agents charge authors for their services, or what the Copyright Clearance Center, which will demand to cover its operating costs, will charge for its services — 37% of all transactions fee, or what even the most famous authors receive in royalties. Added to the fee that the book registry will demand to cover its operating costs, which will probably be around 20%, this means that rightsholders will be getting less than 50% of the income, or not much more than Google itself. I have heard no argument that justifies such a steep toll, and it vastly exceeds the micropayments for advertising upon which Google originally built its multi-billion dollar business. Although Rick Anderson has praised Google because it “has elected to absorb effectively all of the up-front costs and labor involved in this remarkable project,” in fact not a single penny has been provided to pay for the substantial labor costs that publishers will
incurred in researching what digital rights they have in the five million out-of-print works in Google's database, costs that are particularly onerous for small, understaffed university presses like mine to bear. Even finding out what books a publisher can potentially claim in Google's database is not proving easy. Google has provided technical means for searching its database, but so far it is not working very efficiently. Using ISBNs to help a publisher identify its titles, for example, only gets one so far because the ISBN did not come into use until 1970 and in-copyright titles can have publication dates as far back as 1923. One needs to investigate the language in older contracts to see whether it can be interpreted to include any kind of digital rights at all, and commercial publishers have the additional problem of tracking the legal ownership of rights through a long maze of mergers in the publishing business. Looking ahead, publishers must figure out how to handle income deriving from advertising under the settlement, as this has heretofore been a type of revenue that publishers have had to worry about sharing with authors. As one university press director has been quoted as saying, "that's one check I don't want." They also face the daunting prospect of having to enter into negotiations with authors over many rights that the settlement identifies as shared between authors and publishers, such as how much of a book to display. It is easy to understand why this type of negotiation was factored into the settlement: it was, after all, an association of authors who publish trade books and are represented by literary agents that was one of the plaintiffs filing the class-action suit. But this represents only a small, even if influential, segment of the class of authors overall. Academic authors publishing with university presses, for instance, typically transfer all rights in their books to the university before publication because in this sector presses themselves have traditionally taken on the role of serving as literary agents for authors. It imposes a very significant burden on university presses to obligate them to negotiate every right of this kind with their authors, who mostly want to be left alone to pursue their research and are generally not interested, as trade-book authors are, in all the many details of subsidiary rights. The settlement provides no money to press to cover these extra costs. Conceivably, these costs will exceed what income can be expected from "long-tail" sales of out-of-print titles. There is also a strong possibility that, with its makeup evenly divided between representatives of authors and publishers, the book registry will find itself frequently split in the decisions it will have to make, thus leaving it to the prescribed arbitration rules to resolve at least some of the many potential disputes that may arise under the settlement. Lack of control over outcomes is thus another cost that can be anticipated.

Whether the settlement overall will be sufficiently beneficial to make it worthwhile for a publisher to remain in the class instead of opting out altogether and thereby preserving the option of bringing suit later or reaching an agreement with Google outside the terms of the settlement, such as within the alternative framework of the Google Book Search program that already exists, is a complex decision that each publisher will have to make for itself. While the settlement seems a mixed blessing for publishers on the whole, the exact mixture of costs and benefits will vary from one publisher to another depending on a variety of factors different for each, among them the number of titles already in the database that each publisher can credibly lay claim to owning, the degree of complexity anticipated in negotiating the display and other rights with authors, the terms of other agreements a publisher may invoke (such as Google Book Search, if a publisher should decide to bring some now out-of-print titles back in print in such a way as to satisfy the requirement that they be "commercially available"), and the potential monetary rewards under alternative programs compared with the settlement (which guarantees just $60 per title already digitized plus a 63% share, minus the registry’s fee, of income derived from institutional subscriptions according to whatever formula the registry devises) and the likelihood that the terms of alternative agreements outside the settlement will remain relatively favorable upon renewal of those agreements.

There is a great deal of uncertainty right now about how all this new arrangement with Google will work out in the long run — whether, for instance, it will become the veritable pot of gold at the end of the rainbow or, instead, simply income marginal for the publishing industry, which may become a reliable source of extra income but nothing on a scale to revolutionize the business in any fundamental way. Each publisher will be placing its own bets, initially by opting out of or staying in the settlement, and it will be interesting to watch which kind of gamble pays the best returns in the future.

Building Library Collections in the 21st Century
It’s the Economy, People

Column Editor: Arlene Sievers-Hill <axs23@case.edu>

Those of us who have toiled for many years in acquisitions, serials and collection development in academic libraries have met numerous challenges to the budget and the profession. However what we are facing now internationally, nationally in university libraries and personally pale all that came before in our life times.

Building of academic library collections in the 60s, really began on a vast scale, spurred by Cold War politics and an ever increasing number of college students. Approval plans went into effect in the 60s because the building of collections required lots of books and individual purchasing was just not efficient. Subscription agencies also bloomed to manage the increasing number of subscriptions to journals. This was caused by a real decline in university budgets, which in this case may be a lasting impact which may undermine severely our ability upon renewal of those agreements.

In the 60s and 70s there were the serial price wars that waged — American librarians against European STM publishing behemoths. Price differences based on location, and taking into account currency fluctuations, which were difficult to track reliably. Journal prices went up so dramatically that budgets began to be really pinched and book purchasing was reduced due to the serial price increases. Simply not enough money for everything.

Fast forward through the development of the Internet, journal and databases. There was a naivete that existed for a while that we could make as much money as before, even more, as ejournals were sold in packages. Subscription agencies jumped in by grabbing a huge role in the creation of databases which held and indexed these journals. Now, I want to say, as one who worked as both a librarian and in the library subscription industry, I see nothing wrong with these businesses. They operate as businesses to earn money for themselves and their shareholders. Libraries, even ones at expensive universities, are altruistic in principle, and librarians sometimes have a hard time seeing the differences.

Now to the real subject — the kamikaze like death spiral of the world economy and its effect on everything else, including our little world of library collection building and acquisitions. The plunging of securities in the stock markets, the wobbling and failure of national and international banks, and the massive layoffs are all having a current and probably more subsequent and longer lasting impact which may undermine severely what we are trying to do.

The number one effect is the bugabo we have always faced. This is the decline in real dollars to build collections, which in this case may be a caused by a real decline in university budgets, requiring not only priority changes, but cuts everywhere. Already, even before budgets are set for the new financial year in June/July, libraries are cutting back on book approval plans — going from automatic shipments to form only plans.

The impact on the ever-increasing implementation of current plans...