2009

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: http://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: http://dx.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.
outlay. In several organizations that are intermediaries between suppliers and end users, there are large teams of people whose sole job is to clean and append information to publisher-supplied metadata. Obviously, there are significant perceived benefits and a return on the investments for improving the supplied metadata before it is passed on or made available to the broader community. Otherwise these organizations would not invest such significant resources in improving the data.

Improving the interchange of metadata was one of the main recommendations of the Digital Libraries and Digital Collections Thought Leader meeting that NISO sponsored in 2008. The Thought Leader meetings — funded in 2008 by a grant from the Andrew W. Mellon Foundation — were held with the goal of identifying and prioritizing new initiatives of importance to the information community. The group discussing digital collections suggested that NISO sponsor the creation of a suite of tools that publishers could use to assess the quality of the output they are supplying to the community. However, determining the costs and potential savings for publishers of both doing such assessments and improving quality is critical for justifying the investments that likely will be needed. If a compelling case is not made for a return on investment for publishers, it is unlikely that the publishing community would use any compliance tools and even more unlikely that they would invest in any improvements necessary to improve conformance with the various metadata standards.

To address these issues, NISO is co-sponsoring, along with OCLC, some research into the supply chain exchanges including the different needs of the various metadata supply chain stakeholders and the inherent costs. This research will build a map of the supply chain, identifying the hand-offs of metadata between suppliers and users, as well as the requirements that are done with the metadata before further hand-offs, and the costs to the community for transforming metadata. A key component of this project will be the exploration of potential solutions.

OCLC is organizing a by-invitation symposium in March to be hosted at the OCLC offices, that will bring together many of the key participants in the supply chain of metadata in the community. The initial research will be discussed along with the various needs of the organizations exchanging information. We hope that the discussions will identify potential solutions. Among these potential solutions might be an application of OCLC’s Next Generation Cataloging (http://www.oclc.org/partnerships/materials/nextgen/nextgencataloging.html) to which we now have access is, for the first time at least, likely to be the enemy of the good. We've actually made an art out of letting the perfect be the enemy of the good. We could have imagined even five years ago. Is this good? Yes.”

Bernie Sloan, replying to Okerson on December 20, observes: “Sure, people are better off than they were five years ago as far as getting online access to book-based info. And that’s a good thing. I don’t think the critics are necessarily opposed to Google. I think they think the critics are wondering whether the settlement is a step forward or a step back in the journey towards reaching Ann’s goal.” Bonnie Klein worries, in her December 18 message, about the further corrosive effect of the settlement on rights that libraries have traditionally relied upon: “What is at stake are the current exceptions in copyright law — Sections 108, 109, and to a lesser extent 110 — that are key to library operations, whether brick or click. We are moving to accept as common general practice that every instance of online access may be controlled by the copyright owner [or authorized agent] and subject to toll or metered use. Over time this may undermine and erode the relevance and need for Title 17 exceptions.” And Bernie Sloan, on January 14, reminds us of the qualms Siva Vaidhyanathan had initially expressed about the settlement: “My major criticisms of Google Book Search have always come from the perspective of the university libraries that have participated in this program rather than Google itself... Libraries at public universities all over this country... have spent many billions of dollars collecting these books. Now they are just giving away access to one company that is cornering the market on on-line access. They did this without concern for user confidentiality, preservation, image quality, search prowess, metadata standards, or long-term sustainability. They chose the expedient way rather than the best way to build and extend their collections... I am sympathetic to the claim that something is better than nothing and sooner is better than later. But sympathy remains mere sympathy... we must reflect on how complicit some universities have been in centralizing and commercializing knowledge under a single corporate umbrella.”

Others have more explicitly developed Vaidhyanathan’s critique in terms of an alleged monopoly or quasi-monopoly that the settlement has effectively created for Google. Robert Darnton, writing about “Google & the Future of Books” in the New York Review of Books (February 12, 2009), concedes that “Google can make the Enlightenment dream come true,” but reminds us that “the eighteenth-century philosophers saw monopoly as a main obstacle to the diffusion of knowledge — not merely monopolies in general, which stifled trade according to Adam Smith and the Physicocrats, but specific monopolies such as the Stationers’ Company in London and the booksellers’ guild in

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

E veryone 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 libelous listserv on the merits of the settlement. Rick Anderson, in a posting on January 23, prefers to accent 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 aspects” in the settlement and points out: “What I hear from readers is that they are waiting for the day where a click on a library catalog entry will take them directly to the full text of the item and speed up their ability to get information and do research. The Google partnerships and projects bring us closer to a version of that day, much sooner than

Against the Grain / February 2009 <http://www.against-the-grain.com> 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 ineluctable to copyright. Chris Castle, a former lawyer for Napster 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 material, that competitor 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 (MFN) protection. The registry is contractually required to offer Google any better terms it would give to anyone using any data or resources that Google provides the registry, or that is not 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 reward for it, they would be trumped by 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 notes in The Chronicle of Higher Education (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 limns couldn’t yet be trusted with the mission “to protect the historical record and to ensure its future for the public.” 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 future.” 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 rhetoric for universities themselves, had they chosen to do so, to capture this developing market for themselves. But this “missed opportunity” was allowed to pass, and libraries have spent decades now ruining the consequences. In principle, there seems to be no reason that a “grand alliance” of the non-profit kind Darnton limns couldn’t yet be formed to challenge Google commercially, 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 if 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 rhetoric of acting in the public interest always seems to defer to dependence 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 possibilities for selling backlist titles opened up by Google Book Search and its facilitation of “the
long tail,” they were equally dismayed by Google’s preemptive strike against copyright interests in its library digitization program and sided with authors and commercial publishers in their suit, though not formally being a party to it other than their being included in the class of rightsholders once the suit was certified to be a class action. Ambivalent about Google from their past experience, presses seem to have so far accepted the settlement as something of a mixed blessing.

On the one hand, Google’s acceptance of the “opt-out” approach for all in-copyright, in-print books is a major victory for all publishers, as it was mainly to defend this principle against Google’s favored “opt-out” approach that the suit was brought in the first place. Google did get the plaintiffs to agree to the “opt-out” approach for all in-copyright, out-of-print books, and as this category is by far the largest (constituting five of the seven million books already in the Google database, with the remainder equally split between books in the public domain and books still in print and under copyright), Google can boast that in sheer practical terms it won the battle. However, inasmuch as this approach as applied to this category bears considerable similarities to the approach that was embedded in the “orphans works” legislation that both librarians and publishers had been supporting in Congress, it can be considered a reasonable compromise that mostly solves a long-standing problem. (The chief opposition to that legislation has come from creators of images, and it is noteworthy 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 & 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 superseded 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 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 publishers primarily 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 presses 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 into print in such a way as to satisfy the requirements 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 pales all that came before in our life times.

Building of academic library collections in the 60s, really began on a vast scale, spurred by the serial price increases. Simply not enough 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 into 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 pales all that came before in our life times.

Building of academic library collections in the 60s, really began on a vast scale, spurred by the serial price increases. Simply not enough 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 into 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.