

Little Red Herrings — Not with a Bang

by **Mark Y. Herring** (Dean of Library Services, Dacus Library, Winthrop University) <herringm@winthrop.edu>

Numerous stories abound in the news these days, professional and otherwise, references about the **Google Book Search Deal**. Many have weighed in on this, many, I hasten to add, who understand far better than I the far-reaching ramifications of this deal. What I intend to do here is mention some of the more prominent aspects of the deal and (in a second column) end with some reflections on what it may mean for libraries.

In the first case, this is an enormously complicated deal — over 200 pages long when you toss in the appendices — and any one-column-long assessment can only be isagogic. No attempt here is made to try to ferret out all the legalese that only lawyers understand and enjoy. But some items do stand out, and it is to those that we now turn.

Google intends, or rather began with this intention all along, to make all published material — copyrighted or otherwise — available through its search engine. When **Google's** two founders (**Page** and **Brin**) set out and fashioned **BackRub** (**Google's** original name), the enterprising entrepreneurs hoped to create the largest library in the world, bar none, beyond **Alexandria**, beyond **LC**, beyond all of them, combined. **Google Book Search** was the logical next step down that long and winding road, and now the **Google Book Search Deal** is yet another step, or rather a sidestep, to accomplish that first sought-after goal.

Let me emphasize from the outset that I find nothing wrong with the *intent* to create the world's largest library. I would like to own it, and I have no desire to keep knowledge from anyone. But do note the distinction: *knowledge*. Information is one thing and it is everywhere. We are besotted with it. Knowledge is altogether something else, and it requires a great deal more than putting terms in an inquiry search box and pressing a key, though admittedly the latter is most certainly in keeping with our fast-food, sound-bite, instant-messaging world. The trouble is that knowledge is far more expensive and unwieldy than information. Information, by virtue of being easier to harness and given the appearance of being the whole show, is much more enticing and attractive.

Forgive the image, but let me put it this way: information is to knowledge what prostitutes are to sex. The latter may well be attractive, even inviting and, so to say, may well get the job done. Some may argue that they feel as if they really made a “connection” in the liaison. But that is the trouble. They can also be cheap, in more ways than one, tawdry, and, all too often those so-called “connections” end badly, cause scandal, and make one appear, in so many, many

ways, stupid beyond words (I think here of recent politicians). In the final analysis, however, you can't really expect much as far as a relationship is concerned. For that, you must not only look elsewhere but also invest great time, effort and self. So also, it seems to me, this is the difference between information and knowledge. The one is abundant, easy and cheap, the other time-consuming, self-investing and precious. But I digress.

The **Google** deal is on a fast track, to say the least. A hearing has already been scheduled (11 June 2009), *less than one month* following the deadline for filed comments on the settlement. While the settlement today is only a very early excursus, it still represents a positive benefit to **Google**. Whether the same could be said about it with respect to libraries remains to be seen, and I will attempt that question in a second column. A great summary has been written by **Policy Bandwidth's Jonathan Band's** “A Guide for the Perplexed: Libraries and the Google Library Project Settlement” (www.policybandwidth.com/doc/google-settlement-13nov08.pdf). I have read the 200-plus page settlement but have also used **Band's** guide for my own befuddlement.

So far, **Google** has scanned some 7 million books. One million of these are in public domain, another million are in full preview mode. The settlement tries to establish an agreement about the remaining titles and any others that **Google** will add in the future. Under the agreement, **Google** will pay \$125 million to establish a Book Rights Registry to resolve existing claims by authors and publishers, and to cover any legal fees. Suffice it to say that \$125 million is a small price to pay for 7-million-plus titles, but let's leave that for now. Furthermore, \$125 million will be on the order of a class action settlement for authors. The payment they'll receive, save for the most popular among them, will add up to about a penny per one hundred pages. But authors are not the only beneficiaries of this deal. What about libraries? How do they benefit?

The very complicated settlement establishes categories for libraries to gain access to these titles. It does the same for individuals wishing to access the books. For libraries providing content (whether those books are in copyright or in the public domain), one set of rules applies; for others, another. The settlement allows **Google** to continue its scanning and also allows users to search the full contents of the scanned books. The settlement defines three categories of books: commercially available copyrighted books (i.e., those in print or available through print-on-demand); copyrighted books not commercially available; and books in public domain. The settlement has limited reach on the first category as those rightsholders have (and will in all likelihood) control how these books will be used. These books can fall into the default category of the settlement, but it's thought that few authors will likely allow that. The settlement, therefore, applies to the remaining two categories. **Google** estimates that the copyrighted but not commercially available titles comprise 70% of the scanned titles with 20% in public domain. It's

unclear whether this mix will remain when all has been scanned and done.

If you're not terribly confused now, you will be from here on out. All US users will be able to search the entire database of digitized books freely. Public domain titles will yield a full text display. For the copyrighted but not commercially available titles, **Google** will display up to 20% of the text (now only snippets appear). For nonfiction, however, this means only five adjacent pages at a time, or the page you land on and four others adjacent to that page. You can ask for five more where the term is used again, but **Google** will block the two pages before and after any five-page display has already been viewed. For fiction, **Google** displays 15 adjacent pages or 5% (whichever is less), but the 20% cumulative rule applies.

Other rules apply for drama, collections of short stories or poetry, guides, encyclopedias, fiction by multiple authors, quotations and test preparation guides. These rules apply only to those copyrighted books that are not commercially available. For copyrighted materials that are commercially available, users will get only the display of bibliographic information and front matter. On the one hand, the settlement provides for more display than is available now for copyrighted but not commercially available materials; less than what it displays now for copyrighted, commercially available books. Users cannot print or cut and paste from any of these free displays.

Naturally, fee-based services allow for more advantages: cutting, pasting and printing (even of the full text — plagiarism will doubtless increase exponentially, as if it hasn't already), but pricing is governed by “pricing bins.” If users buy the book, they can view it in perpetuity on the site, perpetuity being defined as long as **Google** remains viable. Free Public Access (FPA) is also covered in the settlement and applies to public and university libraries. FPA is defined as terminal access and will likely be limited to one terminal (for colleges and universities this is limited by **Carnegie** category first and per FTE second). Additional terminals, known as Public Access Service (PAS) will be available through an institutional subscription fee but no one knows what this will cost. Users can print from these terminals for a “reasonable” fee but cannot cut and paste. An almost Byzantine security structure (it's 17-pages long in the settlement in a separate appendix) must be in place to prevent unauthorized use outside copyright restrictions and is required of each participating institution. Violations of the security deal range from as little as \$1 (unintentional, single violation) to as much as \$7.5 million (feckless, willful, repeated and intentional).

I've only touched on the most extreme generalities of this “bang” of a deal, but these are enough for now to give readers some idea of the scope and magnitude of it. It should also be obvious by now that many want this deal to succeed. In the next column I will discuss the ramifications of this agreement that will leave some shouting for joy, others whimpering without restraint. 🌸

