## Open Access and Books in a Digital Age
A workshop for the Metropolitan Library System, Chicago, June 9, 2010
Barbara Fister

> When people think of libraries, they think of books, but what does the digital future hold? This workshop will cover mass digitization projects including Google Books, the Open Content Alliance, and the Hathi Trust, will examine trends in book publishing, both trade and scholarly, will consider the impact of the Open Access movement, and will involve attendees in imagining the future of libraries, reading, and books.

**overview of mass digitization projects**

**_Project Gutenberg_** - powered by volunteer labor, has been creating downloadable low-bandwidth text files of public domain books since 1971. Including affiliates, it now has around 100,000 books available with nearly 350 added every month. They are both "*free of charge* **and** *free as in freedom*." Available in low-tech file format so that they can be read on any device.

**_Academic projects_**  **-** From the 1990s, various academic digitization projects have been kicked off such as _The Making of America_ – and others involved with the Digital Library Federation. Also expensive large commercial digital book collections such as _Early English Books Online_ (digital versions of 100,000 books published in English between 1475 and 1700) and _Sabin Americana_ (Joseph Sabin's *Bibliotheca Americana: A Dictionary of Books Relating to America from Its Discovery to the Present Time*).

**_Amazon Search Inside_** – debuted in Fall, 2003. First effort to make it possible to search contents of and read selected sections of current, under-copyright books. The digitization was done with publishers, not with copyright holders. Publishers tend to acquire e-rights (and have argued they own e-rights to books they acquired before they were specified).

**_Google Book Search_** – after being largely shunned by publishers, Google turned to libraries to embark on a massive scanning project that now includes upward of 12 million books. This would incorporate the content of books into their search and challenged the idea of fair use in a digital age, arguing that making a copy in order to index it (and giving the library an archival copy) wasn't infringement. They took on a risk that libraries couldn't afford – and when sued, decided to settle. The problem with this settlement is that the question of fair use never was decided; nobody else can afford to take the risk Google took. Jeffrey Toobin examined what might happen in a 2007 New Yorker article and predicted they would settle – and that it would be an end to competition for digitizing books on a mass scale. Since then, GBS has changed from being a digital library to being a digital bookstore; if the settlement goes through every public library will have one terminal to access the contents of GBS, but academic libraries and libraries that want more than one terminal will have to subscribe; subscription will only allow online viewing, with additional charges for printing pages. For some, the idea that libraries willingly gave away their contents to a corporation is troubling. Others are concerned that GBS may seem like a good reason to pitch old books, but the scanning is not preservation quality. It's interesting to note that Google's annual net revenue is very close to the revenue of all US book publishers represented by the AAP. In many ways, GBS is a nexus for the conflicting goals of publishers and libraries.

**_Open Content Alliance_** – launched in 2005 as a non-profit alternative to GBS, this project aggregates free digital collections including Project Gutenberg and university digitization projects and makes them available through the Internet Archive.

**_Hathi Trust_** – **"**makes the digitized collections of some of the nation's great research libraries available for all." Includes the member libraries' GBS books (those not in the public domain can be searched, but not viewed) and

other digital materials. Willing to play with OCA, but they are not integrated at present.

**Meanwhile, in book publishing . . .**

After a failed romance with e-books around 1999-2000, the launch of Kindle and other dedicated readers has really changed the market for e-books. Kindle has the largest market, but Sony, Kobo, nook, and Alex are other readers on the market; also the multi-purpose iPad and other future tablets; and phones! E-book sales are a tiny part of book publishing revenue, but it's growing fast, and publishers are looking for growth areas. (Sales went from 1.5% to 5% of the market in the first quarter of 2010. Not all reading was on dedicated devices - 37% of eBook consumers were reading on a computer, 32% on the Kindle, 10% on the iPhone, 9% on Sony's reader, 4% on mobile phones, 3% on the Nook and 3% on the iPad—which was on the market for only 3 weeks of this quarter. At this point consumers are not expressing much interest in enhancements – they just want to read. And they are price-sensitive and don't want DRM to get in the way, though they'll accept it if it's not inconvenient.

Here the big issue is that sharing is largely disabled – even the nook, which made a point of allowing some sharing among nook owners, found publishers wouldn't allow it. How can libraries support readers' needs when the texts can't be shared?

Why DRM Doesn't Work (The Brads)

**Differences Between Scholarly and Trade Publishing aka Where Do Books Come From?**

Academic presses see libraries as a significant market; many trade publishers think of libraries as a legitimate piracy scheme. Academic presses are significant brands; prestige is what they represent, and to a large extent they cater to writers who need to be published in order to get tenure. The library market has contracted dramatically, but publishing expectations remain traditional.

Academic presses were once heavily subsidized, but now are expected to cover their costs. University presses are generally very small organizations; some are combining efforts with university libraries, but the coalition is tricky; libraries are focused on the local institution; presses serve the disciplines. Libraries have created institutional repositories, but they are too local and they still depend on traditional publication to vet the contents of IRs. We may see a growing emphasis on open access scholarship, paired with print-on-demand purchase options. The National Academies Press has been offering free online reading for years and considers it a success.

Trade publishers cater to readers, whose connection is with authors, not with publishers, which have no brand recognition. They are distanced from readers by distribution layers and do very little market research. Their leadership still thinks in terms of hardcover sales primarily and is more concerned about piracy than about the positive outcomes of sharing. When they think of e-books, they do not think of libraries.

**The Future of Books**

Some predict that books will be written collaboratively or even remixed endlessly in a liquid library of books without edges or boundaries. Kevin Kelly wrote in his New York Times essay, "Scan This Book!"

> What is the technology telling us? That copies don't count any more. Copies of isolated books, bound between inert covers, soon won't mean much. Copies of their texts, however, will gain in meaning as they multiply by the millions and are flung around the world, indexed and copied again. What counts are the ways in which these common copies of a creative work can be linked, manipulated, annotated, tagged, highlighted, bookmarked, translated, enlivened by other media and sewn together into the universal library. Soon a book outside the library will be like a Web page outside the Web, gasping for air. Indeed, the only way for books to retain their

waning authority in our culture is to wire their texts into the universal library.

Some academic projects invite peer review as the book is written.

[Hacking the Academy](#) - a crowdsourced book composed in one week

[The Book of MPub](#) - a book created by students in a publishing course

Kathleen Fitzpatrick's [Planned Obselescence](#) - a work in progress using commentpress

Several university presses offer free digital editions which they feel do not compete with print sales—but only for select titles, ones that are generally tech oriented (as in Yochai Benkler's [The Wealth of Networks](#)). [Cory Doctorow](#) famously publishes his books with a Creative Commons license and invites remixes. He feels offering online copies free increases his visibility and sales. [Paul Coelho](#) rounded up pirated copies of his books into one Website and saw his sales soar.

But most authors and publishers worry about piracy and are only excited about e-books if they can't be shared; consumers resist DRM—unless it doesn't interfere with their convenience.

Another finding of studies of e-book reading is that contrary to what seems logical, e-book readers are currently better suited to leisure reading, but not for serious study. [Students participating in an experiment](#) at several universities found it frustrating to take notes, highlight text, or flip through a book using a Kindle. They like the interactivity of a book on a laptop—but the eyestrain is frustrating. (Also I have found through interviews that, when writing, they like to arrange texts around them so they can sort and consult many at once.)

A [Highwire survey](#) of librarians in 2009 found that librarians value simplicity and ease of use more than sophisticated features. DRM was the greatest hindrance for patrons. .PDF files are popular – and one commented that it should be like JSTOR. Users want to print and read offline. Publishers generally have a big problem with that. Also, preferences for reading may change as more people use reading devices. A comment: need to be released at the same time as print, priced comparably, and "platform-agnostic." There is also a difference in using online books to search and printed books to read. Having them available in both formats offers different advantages.

Most scholarly books have no particular value unless they are read and shared. Could libraries design a model for publishing quality books, with a print on demand option? Do you think libraries could become publishers and curators of open access books? Several university presses now report to libraries and some are engaging in open access experiments, but the cultures remain distinct and hard to merge. What are the obstacles to libraries becoming university presses?

**Readers and the Reading Experience**

The National Endowment for the Arts published three reports on reading that painted a picture of reading being at imminent and severe risk. The first report had a limited definition of reading—the voluntary reading of fiction—which led to the conclusion that we are facing a massive literacy crisis. This was a great hook for news, and it was widely reported. Much less newsworthy was the third report, which found a slight rebound in the reading of fiction, though reading of books overall was down.

In fact, reading and books remain popular—writing them has *never* been so popular—and

books hold a very particular place in our collective imagination. The dream of digitizing a vast library of books has always been a goal that seems far more daring and desirable than digitizing all television shows or all films. Something about the book itself seems lasting, significant, full of wisdom.

When we surveyed our students we found 93% of them enjoyed reading for pleasure. Though only 3.3% of students said they didn't read because they don't enjoy reading, around 40% of academic librarians surveyed surmised that students don't read for pleasure becuase they don't enjoy it.

Written responses from students in Books & Culture:

In their memories of reading, they empasized closeness, family connnections:

> It was an ordinary place in our house growing up, but it became magical every night when my mom would sink into the soft cushions with a book in her hands. My younger sister and I would sit on either side of her resting our heads against her arms, peering at the illustrations that transformed our living room. My mom's voice would decode the squiggles on the page into words, into a story. My first memory of books comes from this spot in our living room.

> My earliest memory of the library was of story time at my hometown public library. The head librarian's name was Mrs. Pease. She would sit and read to us while we crowded around her in a certain spot in the library to hear a wonderful story told in her animated voice. I remember being very anxious for the children's story time to come . . . I felt very comforted by the readings she read to us because that is what my family members always did with me. I was happy while sitting and being read to.

> My mom and I would curl up on her bed, and she would read [*Little House on the Prairie*] as Laura's life played out in my head. When she was happy, having fun, I was smiling without even realizing it. When she was scared or in trouble, I was bouncing around the bed in a subconscious attempt to relieve the tension. A constant dialogue developed between my mom and I:

> Mom: Are you scared? Do you want to stop?

> Me: No!!! Keep going!

Reading was very much a part of their identity and personal history.

> You know how you hear people talk about how a certain song, a certain food, perhaps a certain smell evokes a memory as strong as if you are re-living it at that very moment? Well, that happens to me, too. Except it's not songs or smells that draw my unsuspecting mind through time. It's books. It is the oddest sensation. I see a book cover, hear an author's name, or maybe just read a word that is prominent in some story or other, and all of a sudden, my mind is being transported back to the time when I was reading that book. I don't just look at a book and see a book. I see my life during that time. I see what I was doing, what I was thinking, what I was feeling. I've always gotten lost in books; I pour an incredible amount of energy into them. Character's problems become my problems. Their delights become my delights . . . They are as much a part of me as my actual memories. Some author out there who has no idea I exist has changed the way I think forever.

> My bookshelf is not just a bookshelf. It's a time warp.

## Can We Preserve Library Values in a World of Digital Books?

How can we protect these values in a digital era? Where do you see problems?

- Access to all
- Privacy in defense of intellectual freedom
- Unbiased representation of a variety of views
- Access to knowledge for personal growth and for the public good
- The "fiction problem" revisisted - how do we respond to the "library as entertainment center" critique?
- Preservation of the past
- The enduring symbolic value of books

Other values?

**Additional reading**

Jonathan Band, GBS March Madness: Paths Forward for the Google Books Settlement
An amazing map of the possibilities that demonstrates the legal complexities.

Jonathan Band, A Guide for the Perplexed Part III: The Amended Settlement Agreement.
A clear guide to implications for libraries of the most recent amendment to the GBS settlement.

Michael Cairns, A Database of Riches: Measuring the Options for Google's Book Settlement Roll Out
Argues that it would be good for Google and for rightsholders to price the library database at a rate that makes it affordable for as many libraries as possible, since it will provide discovery that will be monetized by users who want printouts or copies.

Robert Darnton. "Google and the New Digital Future." *New York Review of Books* December 17, 2009.
Laine Farley and others. "Google & the Future of Books: An Exchange" *New York Review of Books* January 14, 2010.
A premier historian of the book responds critically to the GBS; responses argue the settlement would be good for us.

The Idea of Order: Transforming Research Collections for 21st Century Scholarship. Council on Library and Information Resources, June 2010.
Includes a discussion of the possibility of all-digital library collections, an analysis of the cost of storing a printed book (much higher than digital storage), and an analsyis of a selection of Google-scanned booksreporting concerns about metadata, scan quality, and OCR.

Brewster Kahle. "A Book Grab by Google" *Washington Post* May 19, 2009.
Critical analysis of the proposed settlement and the power it gives Google.

Lawrence Lessig, "For the Love of Culture: Google, Copyright, and Our Future" *The New Republic* January 2010.
Why copyright law stifles culture when subrights prevent sharing of work that includes copyrighted images, films, songs, etc.

The Public Index.
A site for annotating and analyzing the GBS.