

GOOGLE BOOK SEARCH & THE FUTURE OF ACCESS TO BOOKS

Pamela Samuelson
Townsend Center
Feb. 8, 2010

Feb. 8, 2010

Townsend Center

1

GBS IN 2004

- Was not about creating universal digital library to last forever
- G was scanning books from research library collections to index their contents & make short snippets available in response to search queries
 - Provided link to library from which could borrow the books, to bookstores from which could buy
- Research libraries were to get back digital copies of books G copied from their collections
- Google formed partner program through which publishers could allow books to be scanned

Feb. 8, 2010

Townsend Center

2

AUTHORS GUILD v. GOOGLE

- Sept. 2005, 3 AG members sued G for © infringement for scanning books on behalf of a class of all owners of © in books G was scanning from U Michigan library
- 5 trade publishers, including McGraw-Hill, brought a similar suit in Oct.
- G's main defense was fair use: scanning necessary to index, likely to help users find books, maybe more sales would occur
- AAP and AG approached G with idea for new business model for making digital books available
- Settlement of merged class action announced on Oct. 28, 2008
 - With AG for author subclass, AAP for publisher subclass

Feb. 8, 2010

Townsend Center

3

G's VISION OF GBS 2008-09

- Universal digital library of all books
- Free downloads of public domain books
- In-print books available through Google Editions
- OOP books available through institutional subscriptions & consumer purchases
 - G to pay 63% of revenues to Book Rights Registry for distribution to authors & publishers (less its fee)
 - Up to 20% viewable on open Internet searches
- Free public access terminals at public libraries
- Print-disabled to have greater access

Feb. 8, 2010

Townsend Center

4

CORE OF SETTLEMENT

- G to provide \$45M to compensate © owners as to books already scanned (\$60 per book)
- G to fund creation of a new collecting society, the Book Rights Registry, for \$34.5M
- Authors and publishers can sign up with BRR to get payments from that \$45M + share in any new revenues BRR collects that are subject to the revenue split
- 3 initial services:
 - display parts of books in response to user queries (ads run vs. queries)
 - sale of books to individuals (accessible only in the cloud)
 - sale of subscriptions to institutions, such as UC
- G free to scan all books in which US © interest

Feb. 8, 2010

Townsend Center

5

DEFAULT RULES OF GBSS

- G will determine if book is in or out of print
- If in-print, default is no-display of contents
 - RH must opt in to display uses
 - Most in-print RH likely to sign up for GPP
- If OOP, default is display uses OK (including all commercializations)
 - Display of 20% of contents for preview uses
 - Registered RH can opt out, insist on no-display
 - RH can also sign up for GPP for OOP books (and many apparently are)
- RHs can ask for removal of books from corpus
 - But “remove” only means these books are dark-archived
 - Right to remove set to expire in April 2011
- G entitled to make non-display uses of ALL books

Feb. 8, 2010

Townsend Center

6

OBJECTIONS TO GBSS

- @ 400 submissions to Judge Chin on GBSS 1.0 in Sept
 - 60 new ones on GBSS 2.0 in late Jan.
- Some support (@ 40)
 - GBS library partners
 - Civil rights, disabilities groups
 - Some academics who want to do non-consumptive research
- Overwhelming majority object or oppose
 - Antitrust, class action defects
 - States over unclaimed fund provisions
 - France, Germany, foreign RH groups: violates treaties, inadequate notice, anti-© (supposed to get permission first)
 - Authors: anti-©, unfair terms, inadequate compensation
 - Some publishers say unfair burden on them

Feb. 8, 2010

Townsend Center

7

MAIN ACADEMIC AUTHOR OBJ'S

- Risks of price gouging for ISD because of G's de facto monopoly over corpus of millions of OOP books
- Need for backup plan for research library access
- Inadequate privacy protections
- Censorship risks
- Wrong solution to the orphan book problem
- Publisher plaintiffs undermining the deal

Feb. 8, 2010

Townsend Center

8

PRICE GOUGING RISK

- Prices of ISD to be set based on: # of books in the corpus, services provided, & prices of comparable products & services
 - More books, more services = higher prices
 - There are NO comparable products, services, & cannot be unless Congress gives others same right “class” would give G
- Prices may be modest at first to get institutions to subscribe, but history suggests will rise over time to supra-competitive levels, as journal prices have done
- Only check on price-gouging is byzantinely complicated arbitration process in U Michigan side agreement
 - Libraries can complain to UM that prices they are being charged are excessive; UM may initiate arbitration, but will it?
 - No criteria for restraining price, so how can this be effective?

Feb. 8, 2010

Townsend Center

9

TOO BIG TO FAIL

- Let's assume that GBSS is approved in modified form and works as intended for 10-15 years
 - Libraries, researchers may become dependent on it, get rid of books since no longer seem necessary
 - BRR will press for higher prices, G has duty to shareholders
 - Or G could decide to sell the corpus or at least institutional subscriptions to anyone w/o consent
 - G could also go out of business, decide ISD is not a good business
 - What happens then?
 - GBSS envisions that a third party provider could take over, but what if no one is willing?
 - Need for public-regarding backup plan (e.g., library partners can get together and reconstruct the corpus and make it available)
- What if the servers go down?
 - “404 Google Books not found”
 - Hackers may find GBS an attractive target for attacks

Feb. 8, 2010

Townsend Center

10

USER PRIVACY

- Libraries who supplied the books to G have longstanding policies of respecting patron privacy as to books
 - Many librarians would take a bullet (metaphorically) before violating user privacy
- Many provisions in the GBSS call for monitoring of user behaviors, including reading books in the cloud
 - ISD may come with ads, research library to shopping mall?
- G has thus far been unwilling to make significant commitments about respecting GBS user privacy
 - even though willing to do so with other services, such as Google Health
 - GBSS 2.0 says G won't give personal data to BRR w/o legal process
 - Privacy Author Objection gives examples of what G should do

Feb. 8, 2010

Townsend Center

11

CENSORSHIP RISKS

- Google has power under the settlement to remove books from the GBS corpus for editorial or non-editorial reasons
 - Some governments might object to books, ask G to remove; G has bowed to such pressure before
- Google only obliged to make available 85% of the books eligible for the ISD
 - Would be as though the books left out never existed
- Settlement gives Google power to alter the texts of books with permission of the rights holder or the unclaimed work fiduciary

Feb. 8, 2010

Townsend Center

12

ORPHAN BOOKS

- Google will have a monopoly on “orphan books” (i.e., those whose RH cannot be found)
 - Financial Times has estimated that 2.8-5 M of the 32 M U.S. in-© books are orphans
 - Likely to make up big part of ISD
 - GBSS would give G the right to scan all books
 - G able to monetize OOP ones through sales of books “in the cloud,” institutional subscriptions, ads next to displays (unless RH says no)
 - BRR can only license uses of books whose © owners have registered with it
 - Would take act of Congress to open up “orphans” to others

Feb. 8, 2010

Townsend Center

13

ORPHAN FUNDS

- GBSS 1.0 would have allowed funds from unclaimed books to be paid out to BRR-registered rights holders after 5 years
 - Blatant conflict of interest among class members
 - Inconsistent with state unclaimed funds laws
- GBSS 2.0 envisions appointment of unclaimed work “fiduciary” (UWF) to handle this
 - Use funds to find RHs, sign them up
 - After 10 years, pay out \$\$ to literacy charities

Feb. 8, 2010

Townsend Center

14

UWF POWERS

- Can switch default from “no display” to “display” but not the reverse
 - This means unclaimed books could be commercialized even if they are in-print!
 - Need to sign up with BRR to ensure this default isn’t switched
- Can advise about pricing bins, but no control over prices of individual books or ISD pricing
- Can veto discounts, but not recommend them
- Cannot make books available on open access basis

Feb. 8, 2010

Townsend Center

15

WRONG SOLUTION

- Congress, not private parties, should address the orphan book problem
- Inconceivable that Congress would give one company a compulsory license of this breadth
- If RHs can’t be found after 10 years of looking for them, books should either be available for all to use freely or at least be available for licensing by more than G
- GBSS would interfere with legislative prerogatives by setting up escrow regime

Feb. 8, 2010

Townsend Center

16

PUBLISHER EXITS

- G will not make full corpus of 12M (or eventual 50M) books available to the public
 - In-print books not in ISD, available through Google Editions
 - ½ of books in GBS corpus are foreign books from research libraries, most of which are out of GBSS 2.0
 - Many publishers are opting out of the settlement
 - Many, including AAP litigants, excluding OOP books from GBS
 - If settlement is such a good deal, why aren't they staying in?
 - Any rights holder can ask for non-display, removal, or no scan
 - G reserves right to remove up to 15% of books from corpus for editorial or non-editorial reasons
 - Orphan books may end up as the bulk of the corpus
- Universal digital library—"a library to last forever"—unlikely to be achieved through the GBS settlement

Feb. 8, 2010

Townsend Center

17

WHERE THINGS STAND

- Judge Chin is scheduled to hold hearing on GBSS on Feb. 18
- DOJ has told him that they do not believe he has power to approve this settlement
 - Taking occasion of lawsuit on one narrow issue to remake market and restructure rights of © owners; this is a job for Congress
 - AG, AAP inadequacy of representation issues
 - Antitrust problems (e.g., price-fixing, de facto monopoly over corpus of millions of books)
- Judge might disagree, but then will be appealed

Feb. 8, 2010

Townsend Center

18

WHAT WILL GOOGLE DO?

- Probably keep scanning books
- Probably provide snippets in response to searches
- Probably use books to improve search
- Will sell in-print books through Google Editions
- Probably not provide institutional subscriptions to libraries except for books for which it gets permission from rights holders
- May litigate fair use issue if GBSS doesn't get approved (or maybe AAP, AG will drop lawsuit)
- May seek Congressional support for mass digitization of books

Feb. 8, 2010

Townsend Center

19

ALTERNATIVE TO GBS?

- A mass book digitization project akin to the Humane Genome Project should be funded by US, states, foundations, donors
 - Best thing G has done is to show the societal desirability of a searchable corpus of millions of books
- Focus on books in collections of major research university libraries, who should collaborate to make shared research corpus
 - Each will still have unique special collections to differentiate, but collaboration, sharing essential
- Authorize libraries to do the scanning and provide high quality scans & accurate metadata
- Provide snippets unless RH objects
- Protect privacy, license on reasonable terms

Feb. 8, 2010

Townsend Center

20