

*How Research Libraries
Became E-knowledge Networks*

Peter X. Zhou
周欣平

University of California, Berkeley
October 6, 2009

The Trend

- New knowledge is created at an ever faster pace
- Aggregated information and knowledge networks are global
- Research libraries have become evolving e-knowledge networks

Scale

- Large aggregated datasets form the information chain and global knowledge networks
- Open access movement has changed the ways information is disseminated
- Co-existence of print and digital information
- Physical vs. cyber spaces
- Inter-library cooperation leads to global e-knowledge networks

E-knowledge Networks: Characteristics

- Information is dynamic and elastic
- A new economic model, Wikinomics, large scale cooperation and peer-to-peer sharing give rise to rapid production, dissemination, and improvement of knowledge, see **Wikinomics: How Mass Collaboration Changes Everything** by [Tapscott, Don/ Williams, Anthony D.](#)
- The general environment of Web 2.0 and Library 2.0

New Operational Environment

- All library collections on a laptop
- Computing speed at that of the light
- Totally virtual information services
- “Cloud computing” connects numerous servers, creates a large shared cyber infrastructure, and makes gathering and storing information a societal enterprise
- Creation of a single knowledge bank of the humanity

Research Libraries and the E-knowledge Networks

- Libraries are members of the e-knowledge networks
- The demise of the mega-Library
- No library is large enough to meet its users' demand
- Traditional measurements of libraries in collection size, history, and special materials are still important, but need modification

New Considerations

- General capacities and scales of information gathering
- Centrality in the production of knowledge
- Ability to perpetually store and disseminate large quantities of information in many formats
- Innovation
- Open access and resource integration
- Knowledge matrix-- print, non-print, data, etc

A New Type of Library?

- Imperial and private libraries: from Babylon to the French Revolution; in China, from the Han Dynasty to the Republic period
- Public libraries: from the French Revolution to today; in China, from the Republic period to today
- Commercial libraries now begin to emerge in co-existence with public libraries

Three New Models

- Commercial libraries, Google mass digitization & Google Book Settlement
- Non-commercial library network, HathiTrust
- Special cross-institutional research and information infrastructure such as E-science and E-humanities/social sciences initiatives in American Research Libraries in partnership with NSF, Mellon Foundation, etc

Commercial Library Model As Currently Proposed

- Google Print (2004)
- Google Book Search (2005)
- Google Book Settlement (GBS) (2008)
- Google Book Search is not Google Book Settlement
- Goal: To create the largest knowledge/research corpus
- Books for machines as well as humans to read

GBS

■ Collections

- Already digitized 7 million vols, some estimate 15 million vols.
- Public domain: pre-1923 publications
- Current in-right publications
- Out of print, but in right publications

■ Non-consumptive use by research libraries

Scale and Impact

- Total Library Holdings in the U.S. – estimated 40 million titles
- Public domain – 8 million
- In right publications: 32 million
- Out of print 23-25 million
- In right and in trade: 7-9 million
- All world's books can be digitized and searchable, although GBS only concerns U.S. publishers and limited foreign publishers
- OCLC WorldCat titles: 23 millions
- Largest copyright deal in US history, concerning all authors of books

Access Examined

- Free: public domain materials (pre-1923)
- For a fee: “out of print” materials; one free terminal in each public library as currently proposed
- Book Rights Registry (by January 5, 2010)
- “Opt out” option (by April 5, 2011)
- “Alternative registry” and non-U.S. authors
- Different fee structures for individual and institutional subscribers
- Average cost recovery: \$8.69 a book
- Finding books, location identification, purchasing books, advertisement, print on demand

Major Issues and Concerns

- Cartel, exclusive and non-competitive nature of GBS
- The value of “public good”
- Privacy (Google is watching you)
- Data mining and non consumptive use
- Access and cost for libraries
- Metadata
- Information quality: “last scan”? (little possibility for re-scanning on the same scale and scope)
- Our values: libraries are the corner stone of democracy/shared public assets
- Too big to fail

Social Benefits

- Better use of “out of print” books
- Free access to pre-1923 publications
- Creation of the largest library in human history for the public
- Preservation of knowledge
- Unprecedented search capabilities
- Creation of a cultural genome through non-display, non-consumptive use of GBS
- New form(s) of the book, e.g. social model of books, books mixed with other information products

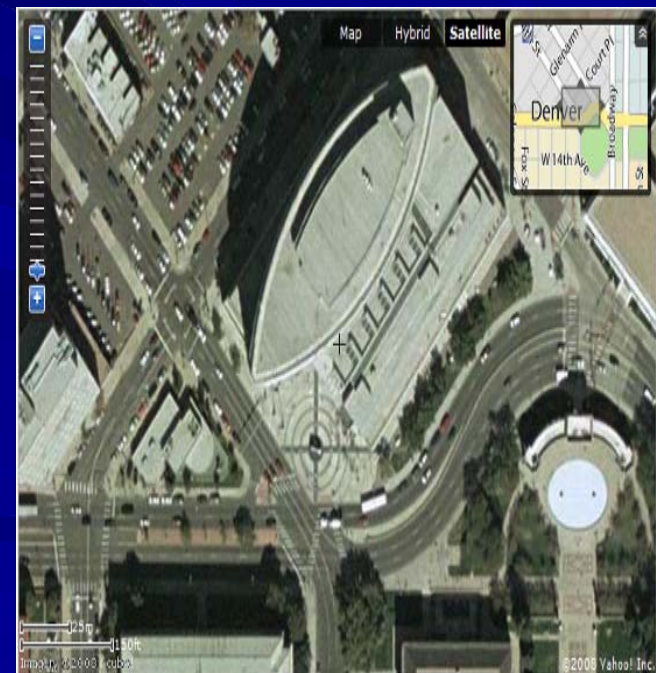
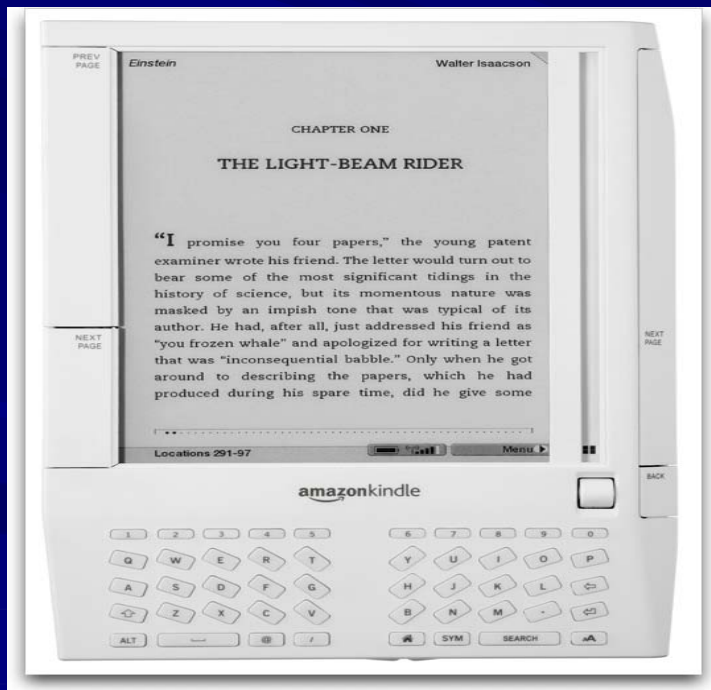
Impact on Libraries

- Total digital library
- Improved access to large corpus of data and collections
- Non-book collections are not part of GBS
- Current book trade
 - Over 90%: print publications
 - 3% e-books
 - Out of print/public domain books take only a small market share

Post-GBS: Challenges and Opportunities

- GBS is not the doom of the library
- Library will not be reduced to an empty hall with only special collections and study spaces
- GBS will change many aspects of library operations
 - OPAC
 - Cataloging practices
 - Interlibrary cooperation
 - Reference and information services
 - Reading service
 - Collection development
 - Preservation

- Our future lies in the alliance of libraries with Google, Internet Archive, Amazon, Open Book Alliance, and other organizations to provide wide e-knowledge networks of content and services



Questions

- What is a library collection?
- What does the library collect?
- How does the library preserve knowledge and what to preserve?
- What must the library give up?
- How do libraries cooperate?
- How to train library personnel?

Thank you!

谢谢！

Research libraries use a very different system, called the Library of Congress (LC) system. Under this system, books are sorted by topic instead of the author. The first section of the LC call number (before the decimal) refers to the subject of the book. Just become familiar with the system. When using a computer, be sure to keep a pencil handy and write down the call number carefully, to avoid sending yourself on a wild goose chase! Remember, it's a good idea to consult the computer and the card catalog, to avoid missing a great source. Developing knowledge base for libraries. requires encapsulation of knowledge into. databases and effectively putting it into. This paper discusses transformational change in academic libraries, as digital technology alters how services are provided, research is conducted, and learning occurs. Highlights include: advantages of libraries over the World Wide Web; redefining the knowledge management paradigm; two different types of information (i.e., explicit and tacit); a vision for online content, access, and services in the new electronic era; the challenges of managing and evaluating Web resources; the need for a search engine that focuses on Web sites, data sets, video clips