Online Catalogues: What Users and Librarians Want

Karen Calhoun
Vice President, WorldCat & Metadata Services, OCLC
None of these challenges can be met by research libraries working independently. They can only be confronted collectively ... Collection development has gone as far as it can go by operating as a set of unilateral city states.”
1. What is the future of library collections and online catalogues?

2. What sorts of materials do scholars, students, and citizens use? How are libraries doing at providing these?

3. User-centered design and “quality” in the user workflow from discovery to delivery

4. What are OCLC’s strategies for helping libraries support discovery and delivery of quality information?
Trends in Librarianship and Libraries

- Pressure on budgets, personnel, and space
- Changing, complex information landscape
- Re-examination of the value of libraries and librarianship
Competition for Resources to Assign to New Initiatives in Libraries

- Engage with institutional or community-based repositories
- Scholarly publishing expertise/communications
- Support for digital asset management in the communities served
- New services for [fill in the blank]
- Develop new alliances, partnerships
- Reveal “hidden collections”
- Integrate library into learning management systems, teaching and research, portals, scholar’s workstation, personal productivity tools
- 24/7 access
- Major space renovation
- Offsite storage
- Next generation systems
A TIME OF TRANSITION FOR LIBRARY COLLECTIONS AND CATALOGUES

Photo: Quite Adept
http://www.flickr.com/photos/quiteadept/4082692761/
What’s the Value of the Print Collections and Collection-Centered Services?

“65% of information requests originate off-campus”
--Discoverability report, p. 4

Data source: ARL Statistics 2007-2008
http://www.arl.org/bm~doc/arlstat08.pdf
What’s the Value of the Print Collections?

$108 million
Renovation of Ohio State University Library:
“The books had come to clutter the library”

http://chronicle.com/blogPost/Library-Renovation-at Ohio/4700
Percentage Change in Median Resources Per Student at ARL Libraries, 2000-2008 (Compared to 2000)

Data source: ARL Statistics 2007-2008
http://www.arl.org/bm~doc/arlstat08.pdf
An Early Earthquake: Where Do You Begin an Online Search for Information on a Topic?

(2005) College Students’ Perceptions of Libraries and Information Resources: a Report to the OCLC Membership:
http://www.oclc.org/reports/perceptionscollege.htm
What Types of Collections Do Catalogues Generally Describe?

Types of Materials Described in the WorldCat Cataloguing Database, 1999-2008
Will the Catalogue Survive In Its Present Form?
“Within the next five years ...

...there will no longer be a monolithic library Web site. Instead library data will be pushed out to many starting places on the Web and directly to users.”

--Provocative Statement #6, Taiga Forum

...we’ll be past the notion that the online catalog is the way you find things in libraries.”

--Interviewee for LC report

http://www.taigaforum.org/documents/ProvocativeStatements.pdf
Revised Google Books Settlement, 14 November 2009

- Includes digitized books from the US, UK, Australia and Canada
- Will create an independent, non-profit Book Rights Registry (BRR) to locate and represent rightholders
- Provides for the sale of institutional subscriptions for online access
WHAT SORTS OF MATERIALS ARE BEING USED? HOW ARE LIBRARIES DOING AT PROVIDING THESE?

The National Science Digital Library (U.S.)
Rising Interest in Digital Collections on the BnF and LC Web Sites

Where do people go on bnf.fr and loc.gov?

**BnF:**
- Expositions: 30%
- Catalogue: 26%
- Gallica: 26%

**LC:**
- American Memory: 41%
- Catalog: 17%
- Legislative information (THOMAS): 6%

Source: Alexa.com, 15 Nov 2009
Open Access Repositories Gaining Visibility and Impact

2008-2009 Traffic Compared:

- Social Science Research Network
- arXiv.org
- Research Papers in Economics
- British Library (bl.uk)

Sources: Alexa.com 15 Nov 2009 and the Cybermetrics Lab’s ranking of top Repositories (disciplinary and institutional) at http://repositories.webometrics.info/about.html
An Increasingly Complex, Demanding Local Environment to Support

- Full Text DBs
- Citation DBs
- E-books
- Institutional repository
- Digital collections
- Online catalog
- Link resolver
- Knowledge Base(s)
- ILS
  - Acquisitions data
  - Circ/status data
  - Print holdings data

E-resource management system (ERM)
The end user perspective: a fragmented, confusing library landscape.

- Full Text DBs
- Citation DBs
- Digital collections
- E-books
- Institutional Repository
- Printed Books & Serials, AV, Maps, Etc.
- Online Catalog Records
- Web Lists
- Online Catalog Records (sometimes)

Diagram: A visual representation of the library landscape with various components like databases, collections, and records connected through arrows.
USER-CENTERED DESIGN AND “QUALITY” IN THE WORKFLOW FROM DISCOVERY TO DELIVERY

Source of image and tag line: Experientia web site
Today's libraries exist in physical and virtual space.
A library is thus both a manifest place and an experience of real, but intangible, “cyberspace” for those who interact with it. One may describe a library system in terms of the relationships between users, collections, library staff, and space, with “space” defined both as buildings and as virtual, networked information space.


Another Type of Space: The Virtual Library (Embedded, on the Web)
<table>
<thead>
<tr>
<th>Trends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users are discovering relevant resources outside library systems</td>
</tr>
<tr>
<td>Users expect discovery and delivery to coincide</td>
</tr>
<tr>
<td>Usage of portable devices is expanding</td>
</tr>
<tr>
<td>Discovery increasingly happens through recommending</td>
</tr>
<tr>
<td>Users increasingly rely on emerging nontraditional information objects</td>
</tr>
</tbody>
</table>
The University of Oxford on iTunes U

Oxford had a double celebration in early August: the millionth download from its iTunes U site and Marianne Talbot occupying the number 1 spot in the iTunes U chart with her "Romp through the history of philosophy".

DISCOVERING RESOURCES OUTSIDE LIBRARY SYSTEMS
Geocentric/Aristotelian view: The local catalog is the sun

Heliocentric/Copernican view: The local catalog is a planet
Data Synchronization and Syndication

WorldCat & WorldCat Partners...

Data synch

BnF

Wikipedia

Facebook

Many other partners
What is Syndication?

For news features like comics, *syndication* publishes the feature in multiple newspapers simultaneously.

*Web syndication* makes website material available to multiple other sites.
WorldCat Partners

Google, Google Books, Google Scholar

ERIC Education Resources Information Center

Greenwood Publishing Group

amazon.com

Biblio

BookPage America’s Book Review

Alchemy

ANTiquarian BOOKSELLERS’ ASSOCIATION OF AMERICA

Celebrating more than 50 years of professional bookselling excellence

http://www.oclc.org/worldcat.org/overview/partnersites/default.htm
WorldCat: Global Integrator, Driving Searches to Libraries

Pushing metadata out, pulling users in: It’s all about linking metadata
Start reading Governing the Commons on your Kindle in under a minute. Don’t have a Kindle? Get your Kindle here.

Governing the Commons. The Evolution of Institutions for Collective Action (Kindle Edition)
by Elinor Ostrom (Author)

Digital List Price: $24.99 What’s this? $
Print List Price: $26.66
Kindle Price: $13.79 includes VAT & international wireless delivery via Amazon Whispernet
You Save: $12.20 (47%)

Text-to-Speech: Enabled ✓

Kindle Books
• Kindle Books include wireless delivery - read your book on your Kindle within a minute of placing your order.

Users expect discovery and delivery to coincide.
“Quality” in the User Workflow from Discovery to Delivery

“A colleague ... sang the praises of the digital world to us. He can now, he told us, get direct access to information ... His enthusiasm had screened out an enormous array of people, organizations, and institutions involved in this “direct” touch. The university, the library, publishers, editors, referees, authors, the computer and infrastructure designers, the cataloguers and library collection managers, right down to the students working their way through college by [working in the library] had no place in his story.”

An End to End View of a High Quality Discovery to Delivery Process

The (invisible) cloud of complexity on the global metadata network

Text
Print
Licensed
Digital
Archival
Data
Images
Sound
Video
Multimedia
Objects
More

Expectation: Easily Find It
AND Easily Get It
Who is using the catalog?

Respondent Type
End Users Only (n=14,895)
Feb. 11 to June 27, 2009

- Student *: 35%
- Teacher or professor: 19%
- Business professional: 18%
- Other (please specify): 12%
- Retired: 9%
- Healthcare professional: 7%
Meet the “user” - primary personas of worldcat.org

<table>
<thead>
<tr>
<th>Doug</th>
<th>Age: 12</th>
<th>GEN Y</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td><em>Uses it</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Call phone</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Free app</em></td>
<td></td>
</tr>
<tr>
<td><strong>Environment</strong></td>
<td><em>Likes it</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>50 miles</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Average</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Has his</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Has it</em></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td><em>Forced</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Soccer</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Goes to</em></td>
<td></td>
</tr>
<tr>
<td><strong>Job</strong></td>
<td><em>Father</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Mom is</em></td>
<td></td>
</tr>
<tr>
<td><strong>Goal</strong></td>
<td><em>Avoid cost</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Get his</em></td>
<td></td>
</tr>
<tr>
<td><strong>Retail/co</strong></td>
<td><em>Likes it</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Uses it</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Wears it</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Not into</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Travels with</em></td>
<td></td>
</tr>
<tr>
<td><strong>Family</strong></td>
<td><em>Expand her</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Be the “best boutique ice cream”</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Want to go national with e-commerce</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Wants to hire people to do the day-to-day administrative tasks (HR, taxes, etc.)</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Wants to have a wedding</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jenni</th>
<th>Age: 30</th>
<th>GEN X</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td><em>Cell phone</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Laptop</em></td>
<td></td>
</tr>
<tr>
<td><strong>Environment</strong></td>
<td><em>Lives in Boulder, Colorado</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Renews an apartment monthly</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Enjoys but doesn’t focus on</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Well-connected in the community</em></td>
<td></td>
</tr>
<tr>
<td><strong>Job</strong></td>
<td><em>Works hard and plays hard, all the time</em></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td><em>Attended University of Colorado</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Got her master’s degree in education</em></td>
<td></td>
</tr>
<tr>
<td><strong>Goal</strong></td>
<td><em>Owns her own business: Jenni’s Ice Cream</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Has one successful shop and restaurant—inclusive flavors</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Summer has been transformed into an interview for school student interns, winter with their own fire</em></td>
<td></td>
</tr>
<tr>
<td><strong>Job scenario</strong></td>
<td><em>Lands a job in Web design</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Sees enough money to re-secure life and re-secure life</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kiet Nitya</th>
<th>Age: 34</th>
<th>GEN X</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td><em>Cell phone</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Laptop</em></td>
<td></td>
</tr>
<tr>
<td><strong>Environment</strong></td>
<td><em>Lives in Atlanta, Georgia</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Emigrated to the US from Thailand</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Career-focused</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Temporarily living with an American who looks for a job</em></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td><em>Has degrees in design and engineering in his home country</em></td>
<td></td>
</tr>
<tr>
<td><strong>Job scenario</strong></td>
<td><em>Wants to land a job in Web design</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Taking online courses to become more marketable</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Writing resume, brushing up on research skills in market</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mari Ananth</th>
<th>Age: 43</th>
<th>LATE BOOMER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td><em>Cell phone</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Laptop</em></td>
<td></td>
</tr>
<tr>
<td><strong>Environment</strong></td>
<td><em>Lives in Silicon Valley, California</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Married with 3 kids: 3, 8, and 13</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Well-appointed house (has cleaning person)</em></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td><em>Lives with her mother</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Has an AA degree in computer science</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Has Certifications (MCSE, C++)</em></td>
<td></td>
</tr>
<tr>
<td><strong>Job scenario</strong></td>
<td><em>Lands a job in Web design</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Sees enough money to re-secure life and re-secure life</em></td>
<td></td>
</tr>
<tr>
<td><strong>Goal</strong></td>
<td><em>Hoping to sponsor family members to come to the US</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Raising two children</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Balance work and family life</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Find time to exercise/do yoga</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Maintain family religious practices</em></td>
<td></td>
</tr>
</tbody>
</table>

**Retail/consumer habits**
- *Wears classic clothes, quality-conscious* |
- *Shopping at Nordstrom* |
- *Plans a nice vacation once a year* |
- *Eats out a lot at nice restaurants* |
- *Vegetarian/vegan at Whole Foods* |
- *Buys the books they need for work* |
- *Listens to books on tape during the commute* |
- *Private tutors as needed* |
- *13-year-old allowance $50 weekly towards gaming* |

**Financial status**
- *Family income of $50K* |
- *Husband also works in industry* |

**Hobbies**
- *Her kids* |
- *Active in Indian and immigrant community* |

**Info sources/influencers**
- *Wall Street Journal* |
- *Technical journals* |
- *Amazon.com* |

**Relationship with library**
- *Uses corporate library* |
- *Summer reading program for children* |
- *Kids do not like to go to the library* |

**It was a good day because...**
- *She negotiated significant discounts on the SAP upgrade from their integration vendor* |
- *Her 10-year-old went the whole day without an accident (booting) training)* |
- *Her husband surprised her with pre-made dinner*
End-Users expect online catalogs:
- to look like popular Web sites
- to have summaries, abstracts, tables of contents
- to link directly to needed information

Librarians expect online catalogs:
- to serve end users’ information needs
- to help staff carry out work responsibilities
- to have accurate, structured data
- to exhibit classical principles of organization

Objectives of our metadata quality research

- Start over without assumptions about what “quality” is
- Identify and compare metadata expectations
  - End users
  - Librarians
- Compare expectations of types of librarians
- Define a new WorldCat quality program...
  - Taking into account the perspectives of all constituencies of WorldCat
    - End users (and subgroups of end users)
    - Librarians (and subgroups of librarians)
What did we learn?

End-user focus group results

Key observations:

• Delivery is as important, if not more important, than discovery.
  • Seamless, easy flow from discovery through delivery is critical.

• Summaries and tables of contents are key elements of a description

• Improved search relevance is necessary.
Recommended Enhancements to WorldCat

End-User Results:
Recommended Enhancements

- More links to online content/full text: 36%
- More subject information: 32%
- More tables of contents: 18%
- Add summaries/abstracts: 18%
- More information in the “details” tab*: 16%
- More author information: 14%
- Add sample text/excerpts: 12%
- Remove duplicate records: 12%
- More edition information: 11%
- More selection of non-English language items: 10%
- More cover art: 10%
- More reader reviews: 9%
- More citation information: 9%
- Add recommendations: 7%
- Add editorial reviews from popular publications: 7%
- More publisher information: 7%
- Increase accuracy (e.g., name, subject headings): 7%
- More format/type information: 6%
- Add ratings: 5%
- Other: 5%

Librarian/Staff Results:
Highlighted Differences

1. Remove duplicate records
2. More edition information
3. More selection of non-English language items
4. More cover art
5. More reader reviews
6. More citation information
7. Add recommendations
8. Add editorial reviews from popular publications
9. More publisher information
10. Increase accuracy (e.g., name, subject headings)
11. More format/type information
12. Add ratings
13. Other
Recommendations from librarian survey

- Merge duplicates
- Make it easier to make corrections to records (fix typos; do upgrades); “social cataloging” experiment—Wikipedia
- More emphasis on accuracy/currency of library holdings
- Enrichment—TOCs, summaries, cover art—work with content suppliers, use APIs, etc.
Research into use and users of digital library collections

“The function of searching across collections is a dream frequently discussed but seldom realized at a robust level. This paper ... discusses how we might move from isolated digital collections to interoperable digital libraries.”

—Howard Besser [4]
Digital libraries >> Digital library aggregators

- Digitized text
  - Books, newspapers, journals ...

- Digitized images
  - Photos, drawings, posters ...
  - Sheet music
  - Maps

- Sound and moving images
  - Recordings, oral histories
  - Film

- Born digital material
  - Scholarly preprints, data sets, dissertations
  - Archived Web sites

*Slovenian government web page, 10 Nov 2002—Internet Archive Wayback machine
Metadata Aggregation for Digital Library Content: Gallica in OAIster in WorldCat

Les poésies du roy de Navarre. Tome 1 / , avec des notes et un glossaire français, précédées de l'histoire des révolutions de la langue française, depuis Charlemagne jusqu'à saint Louis, d'un discours sur l'ancienneté des chansons françaises et de quelques autres pièces...

Author: Thibaud IV (comte de Champagne)
Publisher: 1742 H.-L. Guérin et J. Guérin (Paris)
Edition/Format: Internet resource : French
Database: OAIster
Other Databases: HandPressEbook
Rating: ★★★★★ (not yet rated)  id 0 with reviews : Be the first

More info: http://www.oclc.org/oaister/default.htm
### More Metadata Sources

<table>
<thead>
<tr>
<th>Source Type</th>
<th>Metadata Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professionally produced</td>
<td>- Bibliography - cataloging; abstracting and indexing services</td>
</tr>
<tr>
<td></td>
<td>- Authority and classification data</td>
</tr>
<tr>
<td></td>
<td>- Terminologies</td>
</tr>
<tr>
<td></td>
<td>- Publication supply chain data</td>
</tr>
<tr>
<td>Author/User contributed</td>
<td>- Institutional repositories</td>
</tr>
<tr>
<td></td>
<td>- Scholarly portals (e.g., arXiv.org)</td>
</tr>
<tr>
<td></td>
<td>- Tags, reviews, lists, etc.</td>
</tr>
<tr>
<td>Mined</td>
<td>- WorldCat Identities</td>
</tr>
<tr>
<td></td>
<td>- FRBR Work Sets</td>
</tr>
<tr>
<td></td>
<td>- Facets</td>
</tr>
<tr>
<td></td>
<td>- Full text analysis</td>
</tr>
</tbody>
</table>

Algorithmically produced, reused, harvested...
Lennon, John 1940-1980

Useful Links
- Library of Congress Authority File (English)
- Deutsche Nationalbibliothek Authority File (German)
- Wikipedia John Lennon

Associated Subjects
- Archives
- Assassination
- Beatles
- Biographical films
- Biography
- Biography--Pictorial works
- Chapman, Mark David
- Comedy films
- Criminal psychology
- Criticism, interpretation, etc.
- Documentary films
- Documentary videos
- Drama
- England
- Feature films
- Fiction
- Great Britain
- Harrison, George,--1943-2001
- Historical films
- History
- Interviews
- Juvenile works
- Lennon, John,--1940-1980
- Mass media and peace
- McCartney, Paul
- Motion picture music
- Motion picture music--Excerpts
- Murderers
- Music
- Music videos
- Musical films
- New York (State)
- Nonsense literature
- English
- Notebooks, sketchbooks, etc.
- Ono, Yōko
- Pacifists
- Pictorial works
- Political participation
- Popular culture
- Popular music
- Portraits
- Rock films
- Rock music
- Rock musicians
- Singers
- Social history
- Starr, Ringo
- United States
- United States--Federal Bureau of Investigation
- Video recordings for the hearing impaired
Virtual International Authority File (VIAF)
http://viaf.org/
http://viaf.org/viaf/196844
WHAT USERS AND LIBRARIANS WANT
Everywhere, the Library

Place as Library

Photo: jameshead
http://www.flickr.com/photos/headjames/2665919500/
Bringing writers, readers, and libraries together

- Local catalog linked to a chain of services
- Infrastructure to permit global, national or regional, and local discovery and delivery of information among open, loosely-coupled systems
- Critical mass of digitized publications, special collections, and born digital materials online
- Many starting points on the Web leading to many types of information objects
- Integrate library-managed collections and online spaces for researchers into the user’s workflow on the network
Derring-do


And


Thank You!

Karen Calhoun

calhounk@oclc.org

http://community.oclc.org/metalogue/