Google Amazon: User Searching Behavior And Expectations For Library Catalogs

Athena Salaba, Ph.D.
School of Library & Information Science
Kent State University

Outline

• The user
  - Internet user
  - Library/catalog user
• Technology
• OPAC Improvements
• Trends and issues in cataloging
The Internet user

- PEW Internet studies
  - Search engines
  - User profile

PEW Reports

- 84% of internet users have used search engines
  - 2nd ranking internet activity following email
- Popularity:
  - Google (89.8muv)
  - Yahoo! (68muv),
  - MSN 49.7,
  - Ask Jeeves (43.7muv), and
  - AOL (36.1muv)
- Age groups of internet users:
  - 18-28 42%  51-59 39%
  - 29-40 51%  60-69 31%
  - 41-50 37%  70+ 25%
PEW Reports

• Profile of typical search engine user:
  – Heavy internet user
    • faster connection
    • several times a day
    • is online for long period of times
  – GenX cohort (in their 30s)
  – Socially upscale
    • college educated
    • higher income
    • most likely white than Hispanic than African-American

PEW search engine user study

• Use search engines conservatively – 50% would go back to traditional ways
• Positive about their online search experiences – 92% confident, 87% satisfied
• Naïve about search engines and search results – 62% unaware of the difference between paid & unpaid results
• Use search engines for both important and trivial questions – fact finding, medical information
• Men are more intense and savvy searchers than women.
• Young users are more avid, committed, and trusting searchers than older users.
Search engine users: What do they seek?

• Seasonal information
  - Week-by-week lists of popular searches
  E.g. [http://50.lycos.com/](http://50.lycos.com/)

Google Zeitgeist for week ending 4/17/06

  1. happy easter
  2. proof
  3. boston marathon
  4. bettie page
  5. Irs
  6. pamela rogers
  7. peeps
  8. cindy margolis
  9. mumps
  10. easter egg
  11. aishwarya rai
  12. resurrection
  13. scary movie 4
  14. scientology
  15. tax forms

• More diverse individual interests than before
  - People, places, things
  - Travel, commerce
  - Computers, technology
  - Health information
  - Education
  - Entertainment ... .

• If they don’t use search engines, they use a:
  - Favorite site
  - Familiar portal, database
  - Specific recommended URL
  - Link(s) from one site to another
Information seekers

- Go to search engines first (84%)
- Search for trivial and scholarly information
- Prefer fast and immediate information
- Do not care about accuracy as long as they find something - “good enough”
- Trust both free and purchased resources
- Trust search engines & libraries the same

Information seekers

- Still use the library
  - Not as often, after the Internet
  - Do not foresee to use it less in the future
- Like to self-serve
- Don’t want to learn how to use a system
  - expect it to be simple
- Search engines fit their lifestyle better
Scholarly Information

- More specific, easier to sort through in a database
- Some areas lend themselves to “keyword”
- More and more is becoming available on the Web for free

Library Users

- Not always aware of library’s website and services
- Do not ask a librarian even if a “virtual” librarian is available to them
- See library as a place, have a good idea of the physical requirements
- See library as a place to borrow books
- Reference collections
**Library Users: MORE, FAST, NOW**

- More materials
- More full text resources
- Current information
- Timely
- Integration of all library resources
- Customer orientation

**Catalog Users**

- Easier to search by title & author
- Mostly topical searches
- Keyword preference
- Information literacy
  - Do not know the difference between keyword and subject search
  - Not sure about their search terms
- Do not use the “help” available
- Do not ask for help
- Scope of the catalog
  - Access to what?
Catalog users: Subject

- The most problematic search
- Enter too broad or too narrow terms
- Enter one or two words in the search box
- Do not change their search terms often, if at all
- Repeat identical searches

Catalog users: Subject (LCSH)

- User-system term matching
- Complicated strings
- Understanding subject headings
- Too broad
- Unaware of the syndetic structure, subject browsing – functions of controlled vocabularies
Catalog users: Results list

- Order of results
- Relevance ranking – expected due to search engines
- Choices
- Grouping
- Limiting results

Catalog users: Descriptions

- When a “description” is available, they look in the description and title
- When there is no description, they look in the title and subjects
- Look in subjects for terms to use for searching
- Want content notes and description, summary
- Recognize authors as “authority” in a field
Catalog User: Expectations

- Simple, easy to use
- Have features of popular sites
  - Spell checker
  - Friendly messages:
    - “did you mean?”
    - “more like this”
    - “people interested in this, also liked”
  - Default “AND” operator in keyword search
- Standardized searching techniques
- Retrieve something
- Not have to repeat their search in different systems

Technology

- Google
- Amazon
- Keyword searching dominance?
- OPACs
Technology: Google

Technology: Amazon
Is keyword searching the answer to Google’s popularity?

- Dumping down our catalogs?
- One size does not fit all
- What will users miss?
  - Contextual information
  - Term control
  - Conceptual groupings
  - Conceptual browsing
OPACs

- Intelligent discovery systems
- Make it obvious
  - Easy to use
  - Clear help
- Find all from one place – federated searches
  - Books, media, etc.
  - Journal articles – databases
  - Digital libraries
  - Course materials
- 0 or few hits
  - Spell checker
  - Suggest alternative ways
- Link to full text, the actual resource whenever possible

OPACs (cont.)

- Standardize
- Natural language processing
- Clustering of results
- Offer relevance ranking
- Enrich services: “more like this”
- Enhance records
- Visualization
  - Aquabrowser
OPACs (cont.)

- Recommender feature
- Customization – “My” library catalog
- Hierarchical catalog (FRBR)
- Support non-Roman script searching
- Offer bibliographic services where users are
  - push the catalog to the user
- Ask Vendors to improve user-tasks not
  library tasks
Trends & Issues in Cataloging

- FRBR
- RDA
- Metadata
- Subject access
- Reports
  - Calhoun/LC: Changing Nature of Catalog
  - Mann: Critical review of Calhoun report
  - Marcum: The Future of Cataloging
  - U. of California ....

FRBR

Functional Requirements for Bibliographic Records

- Hierarchical catalog
  - Basic unit of description is the work, emphasis on content instead of the container
  - Group by expression, manifestation
  - Bibliographic relationships
  - Improves collocation
- FictionFinder
- Curiouser
- VTLS
- AustLit
RDA
Resource Description and Access

- Influenced by FRBR
- Principles-based vs. case-based
- Separates content from container (based on type of information, not format)
- Allows for flexibility
  - ISBD punctuation
  - Choice of access points

Metadata

- More diverse resources
  - Cultural heritage
  - Archives
  - Born-digital
  - Research reports
  - Conference websites ....
- MARC
- Other metadata, more appropriate for other types of resources
  - XML based
  - MODS, ONIX, VRA, EAD
- Enriched metadata
Subject Access

- Keyword vs. subject
- LCSH
  - FAST - simpler syntax, faceted
- Browsing
  - Subject
  - Classification
  - Taxonomies
- Collocation: foreign materials

Reports & Papers on future of Cataloging

- Increase of digital resources
- Automatic metadata creation
- Simplified cataloging, less detailed descriptive cataloging
- Streamline cataloging, no need for local policies
- Focus local cataloging on unique collections
- Automatic indexing – relevance ranking – no LCSH?
- Enhance records with vendor/publisher information
What’s next?

- MARC records
- Other metadata
- Use existing descriptive data (publishers, vendors), explore automatic metadata creation
- Focus on special materials, authority control, subject cataloging, and other expert tasks
- Collaborate with other information providers to enhance bibliographic information
  - Cover art, book reviews, excerpts
- Cataloger skills: managers, languages, subject areas, different metadata
- Catalogs that support users
- Catalogs that communicate with external information systems

Questions?

Thank you!
References


