In Search of the Single Search
Box: Building a "First-step"
Library Search Tool

DLF Spring Forum 2005
Tito Sierra and Steven Morris
Problems with Library Search

Systems Perspective

- The non-integrated library system
  - Patchwork of library systems
  - Diversity of interfaces
  - Variety of content formats
- Locally stored versus hosted content
- Metasearch not there yet
Problems with Library Search

User Perspective

- Desire for a single search box that searches everything
- User needs that extend beyond collection specific searches
- Confusion over our library website information architecture (or lack thereof)
- No subject based access integration
Solutions?

- Provide a single search box as a “first-step” in the library search process
- Develop a search tool that responds to a broader range of search needs
- Begin to “interact” with the user in the search space the way we do in reference spaces
- Informally educate our users about our information architecture, tools, and services
Why a Single Search Box?

- Why a search box?
  - Evidence of users ignoring text cues
  - 1/2 of users search dominant, 1/5 link dominant (Jakob Nielsen)
  - “Where is the little box where I can type”

- Why a single search box?
  - “Competing homepage search boxes are confusing, and advanced search should be relegated to a secondary position inside the site to avoid seducing users away from the simple search, which they’re more likely to use correctly.” (Jakob Nielsen)
Search Options: ARL Environmental Scan

Most Prominent Search Option on 113 ARL Library Homepages

- No Search Option: 1%
- Multiple Search Links Only: 29%
- Single Search Box: 52%
- Single Search Link Only: 9%
- Multiple Search Boxes: 9%
Search Options: ARL Environmental Scan

Most Prominent Search Option on 113 ARL Library Homepages

- No Search Option: 1%
- Single Search Box - Catalog: 22%
- Single Search Box - Mixed: 10%
- Single Search Box - Website: 20%
- Multiple Search Links Only: 29%
- Single Search Link Only: 9%
- Multiple Search Boxes: 9%
Search Placement: Peer Analysis
Search Placement: Peer Analysis
Search Placement: Peer Analysis
Search Placement: Peer Analysis

Search Center of Mass:
- Catalog (red)
- Mixed (purple)
- Website (blue)

Catalog Zone

Diagram depicts the placement of search center of mass with catalog, mixed, and website categories.
Eye-track studies suggest upper-left screen area as prime real estate

Current practice: Catalog in prime location, but is it representative of library content/services?

Technically not yet able to develop a comprehensive “upper-left” search tool

Near-term goal: Develop “upper-right” (lower confidence) tool that is more than website search, but less than full metasearch
Library “Quick Search” Tool

- Goals
- Video demo
- Architectural overview
- Description of components
- Challenges
- Future development
Quick Search: film

Are you looking for Media Center (videos, DVDs, etc.)?

Frequently Asked Questions

- Where is the microforms room?
- Where are the Erdahl-Cloyd Theater and the Mini-Theater?
- How can I find videos, DVDs, films, or movies?

Library Webpages

Results 1 - 10 of 110 for film

NCSU Audio-Visual Collection
... alumni, and some film was even rescued from a ...

Memorandum from Arthur Wiltner to Clifford K. Beck
... makes use of film badges and pocket ...
http://www.lib.ncsu.edu/archives/eftext/engineering/reactor/murray/M ...
... 751.html - Cached

Unicorn MFHD fixed field material codes
... Microform Micropaque etc Motion picture Film cartridge mf Motion picture Film ...
http://www.lib.ncsu.edu/cataloging/Unicorn/MFHDmaterialCodes.htm - Cached

LOBO2 - Keyword Builder
... of a word like film and film s. The problem ... be simplified to: ( ...

Search the Collection

Find results for film in:

- Catalog (books & more)
- Quick article search in Academic Search Elite
- Advanced article search
- Journal titles
- E-Reference sources
- Google Scholar (what's this?)

Browse Subjects

- Communication & Media
- Film Studies
- Linguistics
- Literature
- Chemistry
- Materials Science
- Astronomy & Astrophysics
- Mathematics, Science & Technology Education
- Physics
- Computer Science
Quick Search Goals

- Develop a useful “first-step” library search tool that gets users to the right place (without authentication)
- Make it quick and comfortable to use
- Create a platform that allows us to better understand user search behavior
- Contextual presentation of subject based access tools
Video Demo

(7 min)
Quick Search Architecture

- Homegrown application using open source tools (PHP, Nutch, SWISH-E)
- Single query run against multiple specialized indexes
- Locally stored indexes for high performance response
- Modular presentation interface
Quick Search Components in Initial Release

- Library Web Pages
- Frequently Asked Questions
- Search the Collection
- Best Bets
- Smart Subjects
Library Web Pages

- Similar to existing Google site search
- A results set rather than a destination
- No advertising
- Transparent and configurable indexing criteria
Frequently Asked Questions

- Institutional and instructional focus
- Increase visibility and use of existing library FAQ content
- FAQ component displays only if there are keyword matches
Search the Collection

- Research and collection focus
- Stealth instruction about available tools for doing research
- Canned search links
- Low-cost / low-value integration

Search the Collection

Find results for moby dick in:

- Catalog (books & more)
- Quick article search in Academic Search Elite
- Advanced article search
- Journal titles
- E-Reference sources
- Google Scholar (what's this?)
Best Bets

- Featured links to **very** popular content
- Cross functional use
- Responds to alternative spellings and misspellings
- Similar to “sponsored links”, but used as an internal navigation aid
Best Bets Implementation

- Lightweight XML files sources
- Tightly controlled keyword matching to minimize false positives
- Keyword-based retrieval returns the single most relevant Best Bet “ad”, if any

LexisNexis:

```xml
<?xml version="1.0"?>
<ad id="lexisnexis" class="resource">
  <title>LexisNexis Academic</title>
  <keywords>lexis, nexis, lexis nexis, lexis nexus, lexus nexis, lexis-nexis, lexisnexis, lexis, nexus, lexus nexus, lexus-nexus, lexusnexus, lexxis nexxis, lexxus nexus, lexxus nexxus</keywords>
</ad>
```
Best Bets Challenges

- Sensible expansion
  - Criteria for creation of new Best Bet
- Keyword conflict resolution
  - Competing “ads” for similar keywords
  - ‘map’ - library map or map collection?
- Keyword optimization
  - Google AdWords as a model?
Smart Subjects

- Contextual subject links based on user keyword(s)
- An attempt to make new subject based access infrastructure more relevant to the user
- Not really AI, but a contextually filtered view of 96 possible subject nodes
- Subject node landing pages provide subject gateway into metasearch, among other things

Browse Subjects

- Genetics Genomics
- Ecology
- Soil Science
- Entomology
- Plant Biology
- Plant Pathology
- Environmental Science
- Forestry
- Biochemistry
- Chemistry
- Horticulture
- Agriculture

See all subjects
Quick Search: bonds stocks

Library Webpages

Results 1 - 3 of 3 for bonds stocks

Business Information: Network News
... they mean for stocks, bonds, interest rates, currencies... your portfolio ...
http://www.lib.ncsu.edu/news/business - Cached

Business Information: Network News
... they mean for stocks, bonds, interest rates, currencies... your portfolio ...
http://www.lib.ncsu.edu/news/business - Cached

Jack McGieachy's Investment Bookmarks
... Stock Market Chicago Mercantile Exchange Bonds Bonds Online BradyNet: Home Page Discount ... Your Retirement Plan Money.com: Stocks ...
http://www.lib.ncsu.edu/staffjmcgeachy/invest13.htm - Cached

Tip: You may also Search NC State for bonds stocks

Search the Collection

Find results for bonds stocks in:

- Catalog (books & more)
- Quick article search in Academic Search Elite
- Advanced article search
- Journal titles
- E-Reference sources
- Google Scholar (what's this?)

Browse Subjects

- Accounting & Finance
- Business & Management
- Economics
- Agricultural Economics
- Biochemistry
- Chemical Engineering
- Chemistry
- Animal Science
- Biology
### Business & Management

**Related Links:**
- The Business Information Network
- NCSU GIS Data

**Select a Subject Area:**
- Agriculture
- Design
- Education
- Engineering
- Humanities
- Life Sciences
- Management
  - Accounting & Finance
  - Agricultural Economics
  - Economics
- Business & Management
  - Operations Research
  - Public Administration
  - Sports Management
  - Textile & Apparel Management
  - Mathematics
  - Natural Resources
  - Physical Sciences
  - Social Sciences
  - Textiles

**Books on Business & Management**
- Business, Management, Total Quality Management
- More Books...

**Databases for Business & Management**
- EBSOOC Business Source Elite
- Proquest ABI/INFORM Global
- More Databases...

**Selected Journals in Business & Management**
- More Journals...

**Business & Management E-Reference Tools**
- Online Urban Dictionary, Reverse It, The New Treasury, Online Dictionary of Biology, Pennin...
How Contextual Subject Links are Generated

- Subject related content used as index fodder
- Journal article titles, course descriptions as surrogates for subject content
- Index fodder stores organized by campus curricula, or by subject nodes directly
- Simple keyword-based retrieval to return the most relevant subject nodes
- Interesting by-product: institutional bias
Smart Subjects Index Fodder

- Unstructured text dumps of content used as index fodder
- Index fodder generated from multiple sources for subject triangulation and increased keyword coverage
- Shared curriculum code or subject node identifiers tie multiple sources together
- Index fodder stores organized by curriculum require curriculum to subject node crosswalk during retrieval

‘Biochemistry’ (extract from Faculty Pubs DB index fodder source):

1-substituted cyclopropenes: Effective blocking agents for ethylene action in plants Abnormal lignin in a Loblolly pine mutant Absence of association between genetic variation in the promoter of the microsomal triglyceride transfer protein gene and plasma lipoproteins in the Framingham Offspring Study Accurate translation of the genetic code depends on tRNA modified nucleosides Action spectra for UV-light induced RNA-RNA crosslinking in 16S ribosomal RNA in the ribosome Acute

...blah, blah, blah
Smart Subjects Challenges

- Increase opportunities for contextual subject link generation (reduce instances of zero hits)
- Improve relevance ranking of generated subject links
- Optimize the presentation of subject links in the Quick Search environment
Broader Quick Search Challenges

- Maintaining a usable interface as the scope of the tool expands
- Sensible integration
- Keeping Quick Search quick
- Evolving tool to satisfy changing user needs
Future Development

- New search components
  - Embedded journal title matches
  - Embedded catalog results summaries
  - Finding articles / metasearch integration (this is really hard)

- Tool refinement based on observed use
  - Search query log review post-launch
  - Comparative clickthrough stats analysis
Future Development

- Improved index fodder for Smart Subjects
  - Journal TOCs data from Ingenta
  - NCSU ETDs and other campus content
  - Metasearch and/or Google Scholar harvesting?
  - Open Access subject content?
- Make search tool more comfortable to use
  - Smarter “zero hits” response for all components
  - Query expansion for some components (Wordnet)
  - Spelling suggestion feature (Aspell or Google API)
Summary

NCSU Libraries Quick Search

- A “first-step” library search tool
- Our initial attempt towards developing an interactive search space
- A platform for learning about the search needs and behavior of our user community
- Evolving product
- Available Fall 2005, or earlier in beta mode
Questions?

tito_sierra@ncsu.edu
steven_morris@ncsu.edu