Search Present and Future: Implications for Metadata

Carl St. Clair Randall

Presented to CENDI
October 29, 2008

Approved for Public Release
U.S. Government Work (17 USC § 105)
Not copyrighted in the U.S.
Purpose of the Study

• To assess the status of searching methodologies
  – What are the preferred methods of searching and rationale?
  – What are some of the desired search capabilities?
  – What limitations need to be overcome to improve searching?
  – What will be the role of catalogers and indexers in the future?
  – What will searching in the future provide?

* Current Searching Methodology and Retrieval Issues: An Assessment, March 2008 is available from DTIC as ADA478274.
Study Methodology

- Interview questionnaire with 26 questions and statements
- Questionnaire administered in three forms
  - in-person interview
  - e-mail
  - telephone
- Study included 48 information professionals from 29 organizations
  - CENDI member organizations
  - DOD organizations and contractors
  - University information science and computer science departments
  - Information science organizations
  - Other libraries
Selected Study Results

• Preferred Method of Searching
• Limitations of Full Text Searching
• Limitations of Metadata Searching
• Metadata Impact on Search Results
• Improvements Needed in Search and Retrieval
• Future of Searching and Retrieval
Preferred Method of Searching

- Majority of participants used both methods to search
  - Method of searching varied
  - Knowledge of the subject being searched
  - Richness and comprehensiveness of the database
  - Type of information being sought
  - Initial key-word followed by metadata to refine results

- Participants who preferred full-text searching
  - Ease of use and speed
  - Incompleteness and inaccuracy of metadata

- Participants who preferred metadata searching
  - More suited for searching government information
  - Provided better results with good response time and high precession
Limitations of Full Text Searching

- Less relevancy in search results
- Lack of precision and control
- Overwhelming volume of search results
- Lack of synonyms… inability to differentiate meaning
- Limitation in information layout
Limitations of Metadata Searching

- Information seeker at the mercy of the metadata creator
- Inconsistency in metadata presented
- Expensive to create and maintain
- Ambiguity of terms, too few hits and missing categories
- Unfamiliarity with metadata rules may result in poor results
- Requires more education and thought on user’s part
Metadata Impact on Search Results

- More relevant search results
- Provides structure needed for consistency in quality searches
- Can be used to:
  - Narrow and refine search
  - Organize results by category
  - Offer fielded search
  - Display results in a consistent look
- Critical when searching specific collection
Improvements Needed in Search and Retrieval

- More effective search systems
- Ability to search all formats equally -- combine large data sets, multi-media, full text, document component and bibliographic
- Incorporate more secondary sources of information
- Use controlled vocabulary to expand queries and to apply metadata to the records; require metadata for all documents on the web
- Incorporate semantic searching – based on terms and relationships
- Search across domains and better synthesize results
- Address varying learning styles and cognitive processes
- Improve parallel processing architecture and use of distributed processing
Future of Search and Retrieval

• Viewing the future from two approaches
  - How things are going to be
  - How we would like things to be

• Future information seekers
  - Net generation and beyond
  - Expectations of simplicity
  - Desire to find rather than search
  - Information seekers don’t care where information resides
  - Merely want to find information that is needed
Future of Search and Retrieval

- Interactivity and visualization paramount
- Increase capturing of user information
  - Use to modify search engine algorithms and interfaces
  - Improve search results
- Relationships among entities in multi-dimensional forms
- Improvement in human interaction with machine
  - Enable unified access across multiple platforms
- Searching will be universal, pervasive and necessary
- Technical boundaries will disappear
  - Searching available everywhere all the time
  - Consumer and provider increase access to information
- With greater expectations users demands will increase
Future of Search and Retrieval

• Choice of access tools will increase
  - Increase use of web from cellular phones, mobile devices, and television
  - Improvement in bandwidth limitations
  - Television and searching will merge; simultaneous access to broadcast programs and searching

• Powerful tools to combine search functions with data mining operations
  - Allow for search of trends or anomalies in databases
  - Use of visually rich interfaces

• Vertical search engines
Impact on Future of Metadata

• Literature on searching de-emphasizes metadata
• Support user customization
• Metadata to link disparate, multi-dimensional resources
• Needed to incorporate non-text materials
• Increased extraction and automated creation
• Changing role of catalogers and indexers
• Metadata resources may work behind the scenes or be embedded in resources as with semantic technology