

DISCIPLINING SEARCH / SEARCHING DISCIPLINES:

Perspectives from academic communities
on metasearch quality indicators

Katherine Skinner: *Digital Programs Team Leader,
Emory University*

Rohit Chopra: *MetaScholar Fellow,
Woodruff Library, Emory University*

Presentation Overview

- Project context: history, background, and goals
- How do we gather and understand user quality metrics?
- Project goals:
 - to gather foundational data concerning user quality metrics for metasearch retrieval
 - to discover the best way to present retrieval results based on user expectations
 - apply these findings to a working prototype system
 - evaluate the effectiveness of this system relative to standard or typically available alternatives

MetaScholar Initiative

- Digital library research program at Emory University to study applications of new technology for scholarly communication
- \$2.6 M in funding for eight projects 2001-2005
- Collaborations with libraries, archives, museums, and other cultural memory organizations.

MetaScholar Research Areas

□ Metadata Sharing

- Programming services
- Open source tools and systems

□ Search and Discovery Systems

- Combined searching and automated classification
- Crafting and organizing coherent subject domains of related resources

□ Digital Publishing

- Evaluation of new tools for scholarly discourse / content creation

Project Context

- Frustration with current search and discovery systems for scholarly needs
 - Work with Southern Studies faculty group, our Advisory Board
 - AmericanSouth.org, Music of Social Change, MetaCombine

Current Search Systems

- Distributed approach:
 - Journals (ProjectMuse, JSTOR)
 - Books, Movies, Music (EUCLID)
 - Interlibrary Loans (WorldCat)
 - Archival Resources (OAIster, AmericanSouth.org, individual archives)
 - Web-based Resources (Google, Yahoo, etc)

Federated Approach

- Vision for a new, open source toolkit
 - Searching across various informational types
 - Rank according to user profiles and researched “quality metrics” as identified by particular user groups
 - Re-deployable at different institutions for their particular user communities

Staff

- Martin Halbert, Director of Digital Programs and Systems
- Aaron Krowne, Head of Digital Library Research
- Urvashi Gadi, Analyst and Programmer
- Rohit Chopra, MetaScholar Fellow
- Katherine Skinner, Digital Programs Team Leader
- Carrie Finnegan, Graduate Student Associate

Hypothesis

- “End-users of digital library metasearch services think in terms of implicit quality metrics that can be respected in retrieval ranking, improving results over the current text-query relevance and even hub-authority based methods. If effective metasearch systems are to be informed by empirical studies of actual users, this hypothesis must be tested” (Halbert and Krowne 2004)

Methodology

- Qualitative (Emory): Focus groups with Emory faculty and graduate students in the humanities, social sciences and sciences
- Quantitative (Virginia Tech): Survey instrument to record students responses to tasks on metasearch retrieval

Focus Groups Structure

- 9 focus groups between Oct. and Dec. 05
- User communities: Emory faculty and graduate students (in the humanities and sciences)
- 22 participants (11 humanities faculty, 3 sciences faculty, 8 humanities graduate students)

Individual Focus Group Design

- Introduction (objectives, participant rights)
- Hypothetical scenario:
 - Mock-ups of metasearch (for “globalization,” “civil rights movement”)
 - Highlighting metadata attributes (format, no. of citations, substantive quality rankings)
 - Sequence of mock-ups to show increasing levels of metadata / quality metrics
- Comparison with existing metasearch systems (eg., A9.com)

Version 2: Screen a

MetaScholar Search Engine Version 1

Results sorted by **RELEVANCE** AND / OR **DOMAIN** (peer-reviewed journals, books, archival records, websites) AND / OR **NUMBER OF TIMES CITED**

1. Upon this Rock: The Black Church, Civil Rights and the Non-Violence Movement (100 %)

Author: Allison Calhoun-Brown **Source:** Political Science and Politics (Volume 33, no 2, July 2000, pp. 168-174)

Matches: 223 ([see all instances](#)) | **Citations:** 121 articles cite this article ([see all instances](#))

Average rating: ★★★★★

2. Origins of the Civil Rights Movement (100 %)

Author: Aldon D. Morris **Source:** Euclid - Emory Library Database

Matches: 111 ([see all instances](#)) | **Citations:** 255 books cite this book ([see all instances](#))

Average rating: ★★★★★

3. Narrative of the Life of Frederick Douglass: An American Slave (100 %)

Author: Frederick Douglass; Robert O'Meally (Introduction) **Source:** Barnes and Noble

Matches: 245 ([see all instances](#)) | **Citations:** 117 books cite this book ([see all instances](#))

Average rating: ★★★★★

4. The Modern Civil Rights Movement: Federal Records in the National Archives and Records Administration (100 %)

Author: Compiled by Amanda Bylczynski, Glenn Longacre, and Martin Tuohey **Source:** National Archives (NARA)

Matches: 135 ([see all instances](#)) | **Citations:** 235 sources cite this collection ([see all instances](#))

Average rating: ★★

5. Television News and the Civil Rights Struggle: The Views in Virginia and Mississippi (99 %)

Author: William G. Thomas -III **Source:** Southern Spaces (November 3, 2004, <http://southernspaces.org/contents/williamthomas.htm>)

Matches: 3 ([see all instances](#)) | **Citations:** 15 articles cite this article ([see all instances](#))

Average rating: ★★★★★

6. Voting Rights Act ; General Records of the United States Government; Record Group 11; National Archives, August 6 1965; Public law 89-110 (99 %)

Source: National Archives (NARA)

Matches: 2 ([see all instances](#)) | **Citations:** 117 sources cite this collection ([see all instances](#))

Average rating: ★★★★★

7. Abolitionism and Issues of Race and Gender (98 %)

Author: John R. McKivigan **Source:** Library of Congress

Matches: 7 ([see all instances](#)) | **Citations:** 82 books cite this book ([see all instances](#))

















Version 2: Screen b

MetaScholar Search Engine Version 1

civil rights movement

search

Results sorted by RELEVANCE AND / OR DOMAIN (peer-reviewed journals, books, archival records, websites) AND / OR NUMBER OF TIMES CITED

Peer-reviewed journals (see list)  	Books  	Archival Records  
<p>1. Upon This Rock: The Black Church, Nonviolence, and the Civil Rights Movement Allison Calhoun-Brown PS: Political Science and Politics > Vol. 33, No. 2 (Jun., 2000), pp. 168-174 RELEVANCE: 100 percent MATCHES: 223 SEE ALL INSTANCES 124 ARTICLES CITE THIS ARTICLE: SEE ALL Average rating: </p> <hr/> <p>2. Television News and the Civil Rights Struggle: The Views in Virginia and Mississippi William G. Thomas III Southern Spaces > Nov. 3, 2004, http://southernspaces.org/contents/... RELEVANCE: 99 percent MATCHES: 3 SEE ALL INSTANCES 15 ARTICLES CITE THIS ARTICLE SEE ALL Average rating: </p> <hr/> <p>3. Review [untitled] Author of Review: Cecilia A. Moore Reviewed Work(s): Historical Dictionary of the Civil Rights Movement by Ralph E. Luker Church History > Vol. 67, No. 1 (Mar., 1998), pp. 205-206 RELEVANCE: 94 percent MATCHES: 2 SEE ALL INSTANCES SEE 5 MORE REVIEWS OF THIS BOOK Average rating: </p>	<p>1. Origins of the Civil Rights Movements Aldon D. Morris Euclid - Emory Library Collection RELEVANCE: 100 percent MATCHES: 111 SEE ALL INSTANCES 255 BOOKS CITE THIS BOOK SEE ALL Average rating: </p> <hr/> <p>2. Narrative of the Life of Frederick Douglass: An American Slave Frederick Douglass, Robert O'Meally (introduction) Barnes and Noble  RELEVANCE: 100 percent MATCHES: 245 SEE ALL INSTANCES 117 BOOKS CITE THIS BOOK SEE ALL Average rating: </p> <hr/> <p>3. Abolitionism and issues of race and gender John R. McKivigan (ed. and intro.) Library of Congress RELEVANCE: 98 percent MATCHES: 7 SEE ALL INSTANCES 82 BOOKS CITE THIS BOOK SEE ALL Average rating: </p> <hr/> <p>4. The theater is in the street : politics and public performance in sixties America Virginia Tech Library RELEVANCE: 92 percent</p>	<p>1. The Modern Civil Rights Movement: Federal Records in the National Archives and Records Administration Compiled by Amanda Bylczynski, Glenn Longacre, and Martin Tuohy Archive: The National Archives (NARA) RELEVANCE: 100 percent MATCHES: 135 SEE ALL INSTANCES 236 SOURCES CITE THIS COLLECTION SEE ALL Average rating: </p> <hr/> <p>2. Voting Rights Act General Records of the United States Government; Record Group 11; National Archives. August 6, 1965 Public law 89- 110 Archive: The National Archives (NARA) RELEVANCE: 99 percent MATCHES: 2 SEE ALL INSTANCES 117 SOURCES CITE THIS RECORD SEE ALL Average rating: </p> <hr/> <p>3. Photos of the Civil Rights Movement Author: unknown Archive: Atlanta History Center - Kenan Research Ctr. Identifier: oai:atlantahistorycenter:UC400.2JK.1970 RELEVANCE: 90 percent MATCHES: 19 SEE ALL INSTANCES 122 SOURCES CITE THIS RECORD SEE ALL Average rating: </p>

Version 2: Screen c

MetaScholar Search Engine Version 1

Results sorted by RELEVANCE AND / OR DOMAIN (peer-reviewed journals, books, archival records, websites) AND / OR NUMBER OF TIMES CITED

	Peer-reviewed journals <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	Books <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	Archival Records <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> cited more than 200 times	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 44 articles with more than 200 citations 1. Reading more about the Civil Rights Movement... Paula D. McClain (241 citations) PS: Political Science and Politics > Vol. 30... Relevance: 100 percent Matches: 42 (see all) Average Rating: ★★★★★ 2. The Civil Rights Movement's Struggle for Employ... Wayne Arthur Gantoro (273 citations)	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 96 books with more than 200 citations 1. Origins of the Civil Rights Movement Aldon D. Moriss (255 citations) Euclid: Emory Library Collection Relevance: 100 percent Matches: 111 (see all) Average Rating: ★★★★★ 2. The theater is in the street, politics and public... Martin Bradford (219 citations) Virginia Tech Library	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 45 records with more than 200 citations 1. The Modern Civil Rights Movement: Federal... Compiled by Amanda Bylczynski, Glenn Longacre, and Amanda Tuohy (235 citations) The National Archives (NARA) Relevance: 100 percent Matches: 135 (see all) Average Rating: ★★ 2. Music of the Civil Rights Movement LOC research team
	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 89 articles with 100- 200 citations 1. Upon this Rock: The Black Church, Civil Rights... Allison Calhoun-Brown (121 citations) PS: Political Science and Politics > Vol. 33, No. 2 Relevance: 100 percent Matches: 223 (see all) Average Rating: ★★★★★ 2. How the Civil Rights Movement Revitalized Labor... Larry Isaac; Lars Christiansen (99 citations) ASR: American Sociological Review > Vol. 67...	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 134 books with 100- 200 citations 1. Narrative of the Life of Frederick Douglass... Frederick Douglass; Robert O' Meally (intro.) Barnes and Noble (117 citations) Relevance: 100 percent Matches: 245 (see all) Average Rating: ★★★★★ 2. The Brown Decision, Jim Crow and Southern... James C. Cobb (104 citations) Euclid: Emory Library Collection	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 88 records with 100- 200 citations 1. Voting Rights Act General Records of the United States Govt... (117 citations) The National Archives (NARA) Relevance: 99 percent Matches: 2 (see all) Average Rating: ★★★★★ 2. Photos of the Civil Rights Movement Author (unknown) (122 citations)
<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> cited 1 to 100 times	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 219 articles with 1- 100 citations 1. Television News and the Civil Rights Struggle... William G. Thomas III (15 citations) Southern Spaces > Vol 3, November 2004; URL: Relevance: 99 percent Matches: 3 (see all) Average Rating: ★★★★★ 2. A retrospective on the civil rights movement Aldon D. Moriss (54 citations) Annual Review of Sociology > Vol. 25 (1999), pp.	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 134 books with 1- 100 citations 1. Abolitionism and issues of race and gender... John R. McKivigan (ed. and intro.) Library of Congress (82 citations) Relevance: 98 percent Matches: 7 (see all) Average Rating: ★★★★★ 2. Son of the rough south: an uncivil memoir Karl Fleming (67 citations) Georgia Tech. Library	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> See all 134 records with 1- 100 citations 1. A journalist's notes on the civil rights movemen Karl Fleming (88 citations) MARBL: Emory Library Archives Relevance: 100 percent Matches: 106 (see all) Average Rating: ★★★★★ 2. Handwritten Copy of Dr. King Speech, Aug 21 Dr. Martin Luther King Jr. (85 citations) The National Archives (NARA) Relevance: 100 percent Matches: 106 (see all)

Scope of Discussion

- Quality and ‘fitness’ metrics (substantive and descriptive attributes)
- Functionality / logic of retrieval and display
- Aesthetic / design issues

Findings

- Strong affirmation of need for good metadata:
 - at discipline and topical specialization level
 - at genre / domain level (archives, pamphlets, books, etc.)
- Understanding of quality / fitness contingent more on metadisciplinary and disciplinary orientation (sciences, humanities) than on research level (faculty or graduate students)
- Partial consensus across communities on abstract notions of academic quality
- Highly specific information requirements about quality/fitness of records, based on combination of task, discipline, and topical specialization

Demand for Transparency

- Transparency: Explanation of logic of ranking, meaning of each attribute (domain, format, number of citations), range of databases, sources searched etc.
- Dissatisfaction expressed about opacity of ranking criteria in current systems
- Transparency needs to be reinforced through institutional credibility

Empowerment to Evaluate Quality

- ❑ Non-quantifiability of academic value: Deep skepticism expressed about substantive quality indicators such as points system for ranking books and articles
- ❑ Quality defined as the empowerment of academic communities to make their own judgment about quality
- ❑ A ranking system for academic websites described as useful, if provided by legitimate associations eg., the AHA for history websites.

Competing Notions of Quality

- Quality ranking requires balancing competing or contradictory imperatives
- need to find “an equivalent of the academic peer-review process” in terms of credibility
- need to avoid reinforcing “the very conservative base of peer review” while providing access to the broadest range of source materials for research

Discipline- and Task-specific Metrics

- ❑ Theology: divide records by “secular” and “theological” domains
- ❑ Interdisciplinary Studies: metadata on materials such as pamphlets and posters
- ❑ Mathematics: Search intelligence to be sensitive to mathematical terminology
- ❑ Biology: Identification of laboratory protocols in preliminary research for new methodology
- ❑ History, Literature: Archival metadata, full content of archives, and quality of archival holdings

Increased User Options as Quality

- Option to sort full text from summaries etc.
- Option to retain ‘memory’ of search sequence, if trying new options
- Concatenate results of same source from multiple access points (eg., one periodical accessible from multiple databases)
- Option to sort results by word length, number of citations (eg., Amazon.com)
- Option to switch sets of metrics midway through search

Areas for Developing Quality Metrics

- Digital images
- Archives: collection-level and item-level holdings
- Foreign-language content
- Pamphlets, posters, leaflets: cultural resources
- Websites (not counting online journal databases online etc.)

Conclusion: QM Project Phase II

- Creation of prototype system
- Specialized focus groups
 - Archives and archival holdings
 - Foreign language content
 - ‘Non-standard’ academic materials
- Reports