Faceted Navigation and Browsing Features in New OPACs: Robust Support for Scholarly Information Seeking?

Kathryn La Barre

The Graduate School of Library and Information Science,
University of Illinois at Urbana-Champaign, 501 E. Daniel Street,
MC-493, Champaign, IL 61820-6211 USA, klabarre@uiuc.edu

Kathryn La Barre has a Ph.D. in Information Science from Indiana University (2006), and is an Assistant Professor at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Her research interests include: information organization and access systems and structures; interactions between theoretical approaches in information organization and access and current practices in digital environments; and the historical and theoretical foundations of library and information science.

La Barre, Kathryn. **Faceted Navigation and Browsing Features in New OPACs: A More Robust Solution to Problems of Information Seekers?** *Knowledge Organization, 34(2)*, 78-90. 65 references.

ABSTRACT: At the end of 2005, impending digitization efforts and several developments related to the creation of access and discovery tools for informational and cultural objects resulted in a series of responses that continue to ripple throughout the library, museum and archive communities. These developments have broad implications for all three communities because of the goals shared by each in the creation of description, control and enhanced access to informational and cultural objects. This position paper will consider new implementations of faceted navigation and browsing features in online catalogs. It is also a response to challenges to develop interwoven approaches to the study of information seeking and the design and implementation of search and discovery systems. Urgently needed during this time of experimentation, development and implementation is a framework for system evaluation and critical analysis of needed and missing features that is grounded in traditional principles, borne out by practice. Such a framework could extend feature analysis protocols established during the early years of online catalog development.

1. Overview

In November, 2005, James Billington, the Librarian of Congress, proposed the creation of a "World Digital Library" of manuscripts and multimedia materials in order to "bring together online, rare and unique cultural materials." Google became the first private sector partner for this project with a pledge of 3 million dollars (http://www.loc.gov/today/pr/2005/05-250.html). One month later, the Bibliographic Services Task Force of the University of California Libraries released a report (2005): *Rethinking how we provide bibliographic services for the University of California*. Key proposals included the necessity of enhancing search and retrieval, redesigning the library catalog or OPAC (Online Public Access Catalog), encouraging the adoption of new cataloging practices, and supporting continuous improvements to digital access. By mid-January, 2006, the tenor of discussion reached fever pitch. On January 12, 2006, the North Carolina State University (NCSU) Library announced the deployment of a revolutionary implementation for their OPAC of *Endeca's ProFind*™, which until now had only been used in commercial e-commerce or other business applications. NCSU made the bold claim that "the

speed and flexibility of popular online search engines" had now entered the world of the online catalog through the use of faceted navigation and browsing (NCSU, online).

A few days later, Indiana University posted *A White Paper on the Future of Cataloging at Indiana University* (Byrd et al. 2006), which served to identify current trends with direct impact on cataloging operations and defined possible new roles for the online catalog and cataloging staff at Indiana University. The Indiana report was a response to an earlier discussion regarding *The Future of Cataloging* put forth by Deanna Marcum (2005), Associate Librarian for Library Services at the Library of Congress. Marcum posed a provocative series of questions and assertions based in part on the Pew Internet and American Life Project study: *Counting on the Internet* (Horrigan and Rainey 2002): "Do we need to provide detailed cataloging information for digitized materials? Or can we think of Google as the catalog?"

Following Marcum's comments, and the announcement of the "World Digital Library," the Library of Congress released a commissioned report in March 2006, *The changing nature of the catalog and its integration with other discovery tools* (Calhoun 2006). This report contained blueprints for change to Library of Congress cataloging processes, advocated integration of the catalog with other discovery tools, included suggestions that the *Library of Congress Subject Headings* (*LCSH*), long used to support subject access to a variety of cultural objects, be dismantled while arguing that fast access to materials should replace the current standard of full bibliographic records. These arguments were supported by assertions that users seem to prefer the ease of Google over the catalog, and that the proposed changes would place the Library of Congress in a better market position to provide users with the services they want most (Fast and Campbell 2004; OCLC 2002). The ensuing debates served to crystallize the intersection and convergence of the traditional missions of the library, archive ,and museum (LAM) communities to describe, and provide enhanced access to informational and cultural objects. One consistent theme emerged: What competencies and roles can each community bring to bear upon discussions of digitization, access and discovery, in order to provide solutions for user needs?

The library community had a ready answer. Originally designed to provide inventory, acquisitions, and circulation support for library staff, the modern library catalog was designed according to a set of principles and objectives as described by Charles Ammi Cutter in 1876. These principles and objectives underpin the core competency of the library community to create bibliographic records designed to assist users in the following tasks: to find (by author, title and subject), and to identify, select and obtain material that is of interest to them. Discussions about the aims of the catalog are not new and have been ongoing since the early 1970s when the earliest forays of the catalog into the digital age began (Cochrane, 1978). The role played by metadata (i.e. bibliographic records assembled in catalogs), as well as the central importance of search and retrieval mechanisms have long been central players in proposed solutions to providing better services to users. Thus, the suggestions of staff at the Library of Congress, that digitization is tantamount to access, and that search engines, like Google, may supplant the catalog as the chief means of access to cultural and informational materials, have galvanized action throughout the library and information science community. It is critical that any discussions and recommended solutions maintain a holistic view of the principles and objectives of the catalog.

The actions and continuing discussions drew heavily from several sources, including the experiences of the LAM communities with the creation of metadata standards, making data 'work harder' and be more accessible, Web 2.0 applications, folksonomy and social classification, and the importance of leveraging rather than abandoning legacy access systems in a time of spiraling costs and decreasing budgets. For archived discussions of these issues see: NGC4LIB listserv (Next Generation Catalogs for Libraries http://listserv.nd.edu/archives/ngc4lib.html) and Web4LIB discussion list (http://lists.webjunction.org/web4lib/). Another valuable source is Lorcan Dempsey's blog, *Of libraries, services and networks* (http://orweblog.oclc.org/).

To better leverage some legacy subject access systems, such as bibliographic records encoded with MARC, it has been proposed that more (not less) should be done to process these records and corresponding authority files (e.g. thesauri and other controlled vocabularies) in a manner that allows optimal access through the faceted navigation and browsing features of new search and discovery systems (Anderson and Pérez-Carballo 2005, Dempsey 2005). Careful consideration of the information tasks and strategies of librarians, researchers and scholars as they interact with the newly implemented systems, is a critical step in supporting the creation of optimal discov-

ery tools such as search engines, web portals, bibliographic database websites and online catalogs. For the purpose of this discussion and proposed research agenda, users are broadly conceived as library faculty and staff as well as academic faculty and students. In order to further scaffold discussion about such an approach, a research proposal will be described in brief that would seek to develop an integrated conceptual framework for the design of information access and discovery systems for scholarly users. The present position paper is also informed by several previous studies and recommendations. The first is a study of faceted browsing and navigation in websites that used a wireframe-based content analytical approach in order to uncover potentially useful search and discovery features (La Barre 2006). Kuhlthau (2005), Saracevic (1997), and others have long recommended integrated studies of information seeking and information retrieval systems in order to build conceptual frameworks that can enhance proposed information discovery and access solutions. Next, several recent studies of the information needs of scholars and researchers, which lay important groundwork for such an approach, will be discussed briefly.

2. Related research studies

2.1 JISC (2006), RIN (2007), DLF/CLIR (2002)

Two reports recently released in the United Kingdom seek to address the information needs of disciplinary scholars and professional researchers in the UK by identifying uses of library search and discovery services. Each study is viewed as a preliminary step to recommending solutions for optimal search and discovery system features. The first report (Sparks 2005) was sponsored by JISC (Joint Information Systems Committee) Scholarly Communications Group in support of Information and Communications Technologies (ICTs) for education and research. Sparks examined the effect of disciplinary differences on the use of information resources by 780 UK research academics at a variety of institutions and departments. The second study, released in November 2006, was sponsored by the Research Information Network (RIN) as part of an agenda to "promote better arrangements for researchers to find out what information resources relevant to their work are available, where these are, and how they may have access to them." Telephone interviews were conducted with 400 scholarly and professional researchers and 50 librarians (Research Information Network 2006).

Both reports directly address some of the same issues as the 2002 Digital Library Federation/Council on Library and Information Resources (DLF/CLIR) study (Friedlander 2002) which surveyed 3,234 faculty members, graduate students, and undergraduate students in order to collect data on the value, use and perceptions of library services. While the DLF/CLIR study reaches beyond either of the UK studies, some of the data regarding user activities and perceptions are analyzed by discipline, allowing comparisons across all three studies. Other highly regarded studies of information seeking are also central to the proposed study, including Tenopir's (2003) survey of 200 major studies of use of electronic sources by all types of user groups, and Unsworth's (2000) discussion of "scholarly primitives" in common use by humanities and other groups of scholars. Also useful is research that examines the information seeking behavior of researchers in specific disciplines such as social scientists (Meho and Tibbo 2003), humanities scholars (Bates 1995, Palmer and Neumann 2002) and scientists (Murphy 2003, Palmer 1996). General models of information seeking behavior will also be useful to the construction of feature analysis protocols (Cochrane 1981, Wilson 2000).

2.2 Hildreth (1982)

> The current state-of-the-art in the design of online catalogs can be characterized as successful confusion. Much has been accomplished overnight by isolated design teams. Some have achieved superb simplicity, others have provided awesome search and retrieval power (Hildreth 1982, 37).

This could have been written about many of the online catalogs implemented during the past year. Hildreth's 1995 report, *Online Catalog Design Models: Are We Moving in the Right Direction?* proposed a number of next generation catalog features, but remarkably few of these features were common in what Hildreth termed the second generation catalogs of the 1980s and 1990s:

1. Natural Language Query Expressions (in your own language, what it is you are looking for?);
2. Automatic Term Conversion/Matching Aids (spelling correction, Soundex, intelligent stemming, synonym tables, etc.);
3. Closest, Best-Match Retrieval (unlike Boolean queries, doesn't require exact match to be re-

trieved as possibly relevant; matching documents are weighted for ranking);
4. Ranked Retrieval Output (many ranking criteria: most likely to be relevant first, most recent, most cited, most circulated, etc.);
5. Relevance Feedback Methods ("give me more like this one;" "what else do you have on this topic?;" "this book is not at all what I want!");
6. Hypertext, Related-Record Searching & Browsing;
7. Integration of Keyword, Controlled Vocabulary, and Classification-Based Search Approaches; and,
8. Expanded Coverage and Scope (the "full-collection access tool").

In this report, Hildreth encouraged innovative design of "adaptable, adaptive but also collaborative" interfaces for multiple users and a reconceptualization of the "developing" user over time. By calling for a more nuanced view of the user across multiple dimensions, by task, experience and role, he drew increased attention to the importance of communication. His work also consolidated understanding that the "user interface" should be broadly conceived to include the physical components of the work area, and the setting and staff of the organization that hosts the interface (Hildreth 1982, 43). The critical analysis framework provided by Hildreth (1982, 1995) will be the basis for the evaluative process proposed here.

2.3 La Barre (2006)

La Barre (2006) was an exploratory study into the use of facet analysis and faceted classification in website construction and design. Interviews were conducted with 18 information architects and knowledge managers who self-identified as creators of faceted systems for search and navigation on websites. A stratified random sample of 200 websites was selected and subjected to content analysis to find evidence of the use of facets or facet analysis. By drawing on the findings of this earlier exploratory study and upon the research tools, and coding manuals created for it, the proposed study will utilize a new method of capturing web content (La Barre 2006) by the creation of wireframes (see Figures 1 and 2).

Figure 1 is a screen capture of the Department of Labor homepage. It is a colorful and information rich page. Figure 2, below, is a wireframe that has been created of the same website. Wireframes allow site designers to communicate with a design team about content placement and other issues. Here, the wireframe is not generated by the system designer for webpage design, rather it is created by the researcher in a process of deconstruction. Note in Figure 2 how this approach de-saturates the webpage. The wireframe contains features that support search

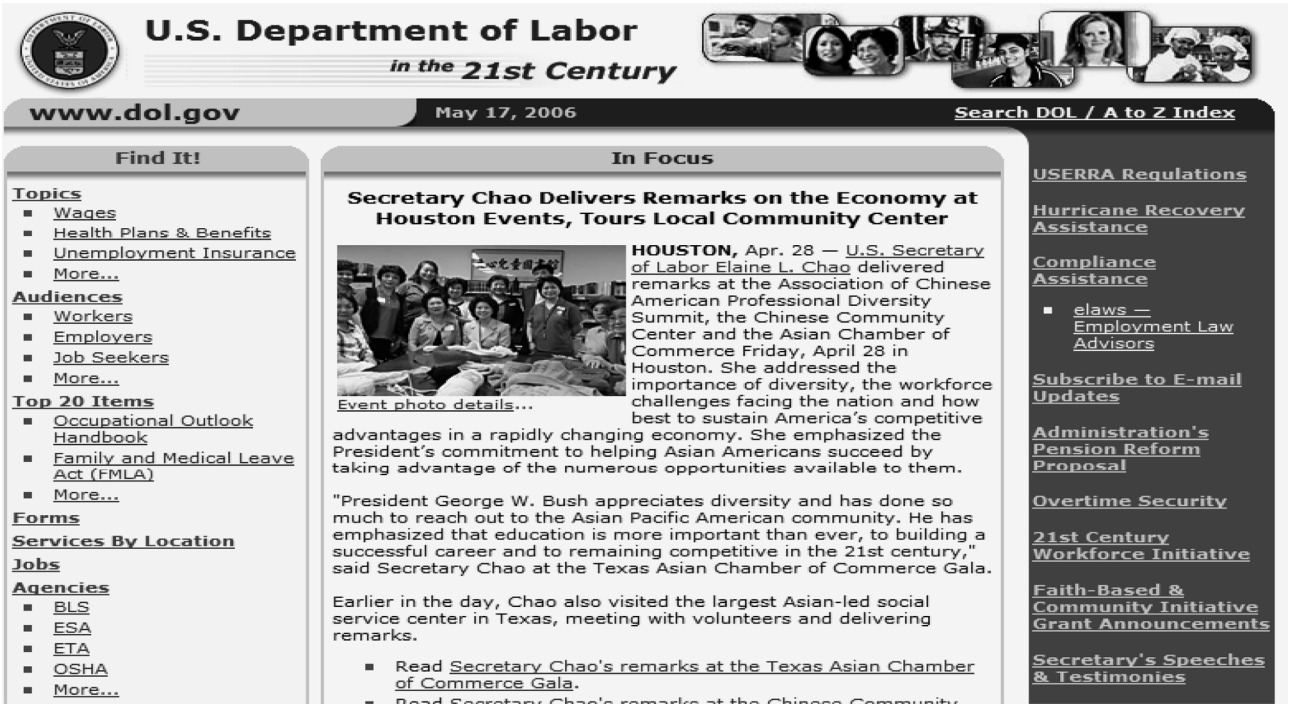

Figure 1: *Department of Labor homepage. http://www.dol.gov*

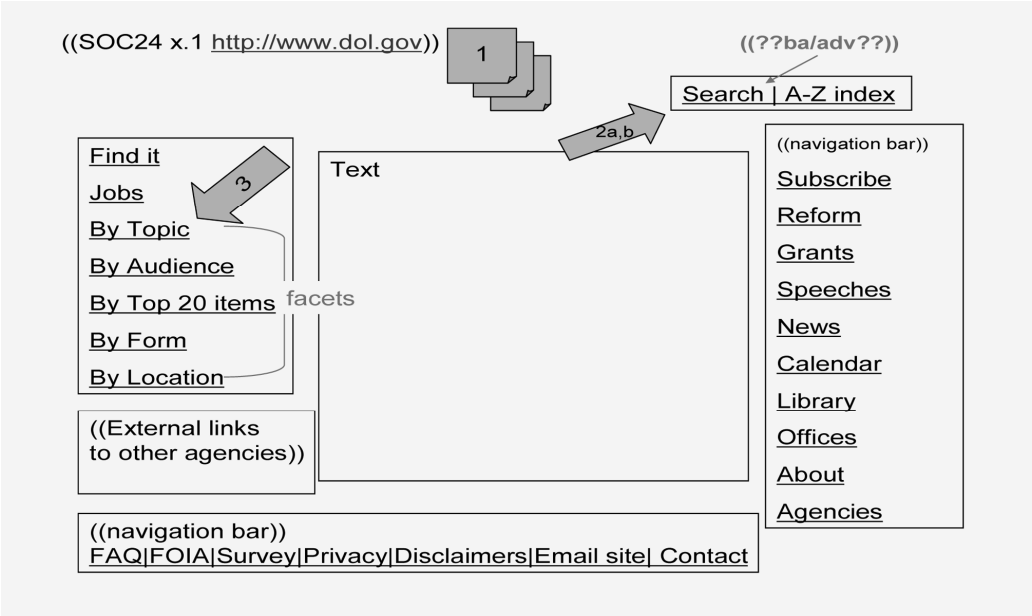

Figure 2. *Wireframe of Department of Labor homepage demonstrating facets (or categories) that assist users in site navigation and browsing*

and navigation, such as hyperlinks to other areas of the site while eliminating extraneous text and other images. The full text of the page is not captured, as it is not the focus of analysis.

The wireframe approach allows the features and position of elements of interest on each webpage to be captured for later viewing, and further analysis. In Figure 2, the wireframe has been coded to indicate the navigation bars and the basic and advanced search mechanisms on the site. "Facets" (as defined by the interviewees in the 2006 study) are also evident on this page on the left hand side. A more traditional understanding of facets is that facets represent the categories, properties, attributes, characteristics, relations, functions or concepts that are central to the set of documents or entities being organized and which are of particular interest to the user group for whom a particular resource is being created. On the Department of Labor website in Figures 1 and 2, the same facets that appear on the homepage: topic, audience, form and location; are used throughout the site (which has three levels as indicated by the coding in Figure 2) as ways to help users search and navigate to find information, literature, forms and office locations. One finding from La Barre (2006) is particularly instructive. The participants in this study described the systems they designed and implemented as faceted systems, yet were largely unaware of the principles of facet analysis and faceted classification. Thus suggestions for system improvements, drawn from theory, were made in order to augment design practice.

While previous information-seeking studies sought to map and quantify scholarly use of different sources, identify areas of difficulty in obtaining sources, and determine perceived values for library services, these studies often focus broadly on the whole of scholarly communication, or narrowly focus on only the electronic transmission of scholarly information. The intention of the proposed research is to focus directly on the interactions between information seeker and the discovery and access tools themselves by examining new interfaces that claim to utilize 'faceted navigation and browsing,' such as the NCSU OPAC implementation of *Endeca* (http://www.lib.ncsu.edu/endeca/). Other examples pictured here are the Nelsonville Public Library's implementation of the open source™ software *kohaZOOM* (http://search.athenscounty.lib.oh.us/) and the Queens Library implementation of *Aquabrowser* (http://aqua.queenslibrary.org/). Figure 3, below, demonstrates the opening screen, while Figure 4 shows the use of "facets" as presented to the user in order to allow for search refinement, or browsing of the result set.

The following screenshots in Figures 5 and 6 demonstrate the implementation of *Aquabrowser* in the Queen's Library OPAC. On the left side is the dynamic display of terms drawn from the result set and displayed by Aquabrowser. On the right side of

Figure 3. *Opening screen of the Nelsonville Public Library – using the open source software kohaZOOM. (Screen capture, June 2007).*

Figure 4. *Search result screen at the Nelsonville Public Library – showing facets for refinement of the search query, or for browsing the result set. (Screen capture, June 2007).*

the screen is the set of facets for search refinement or browsing.

Another essential part of the critical feature analysis will be an examination of the validity of the 'facets' used by these interfaces in comparison with a set of facets generated through the method of facet analysis as described by traditional, or canonical literature about facets and faceted classification (La Barre 2006, Ranganathan 1962). Table 1 contains a set of exemplars that followed the canonical approach prior to system design and implementation, and can provide guidance for the comparative part of this study. Of special note is the work of Pollitt (et al. 1994, 1996, 1998), who created early online prototypes that implemented faceted views in databases and OPACs.

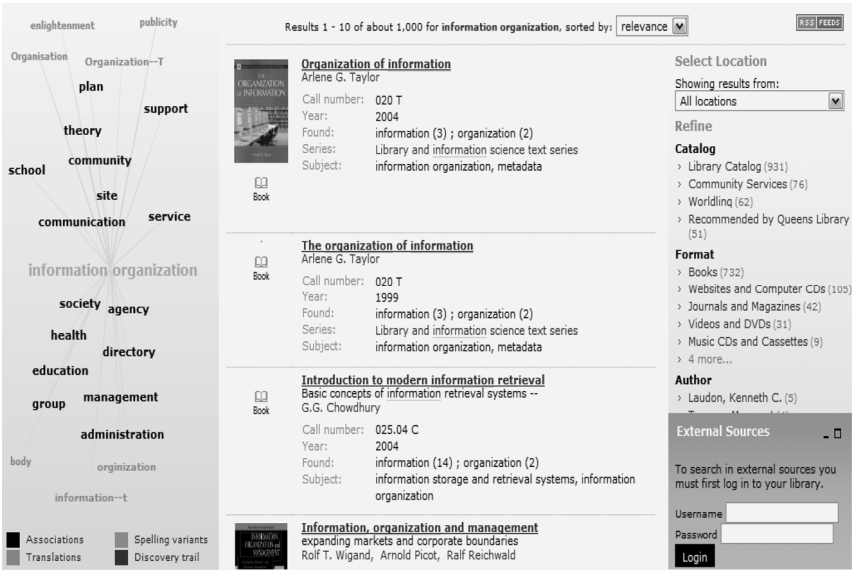

Figure 5. *Search result screen from the Queens Library implementation of Aquabrowser. (Screen capture, June 2007).*

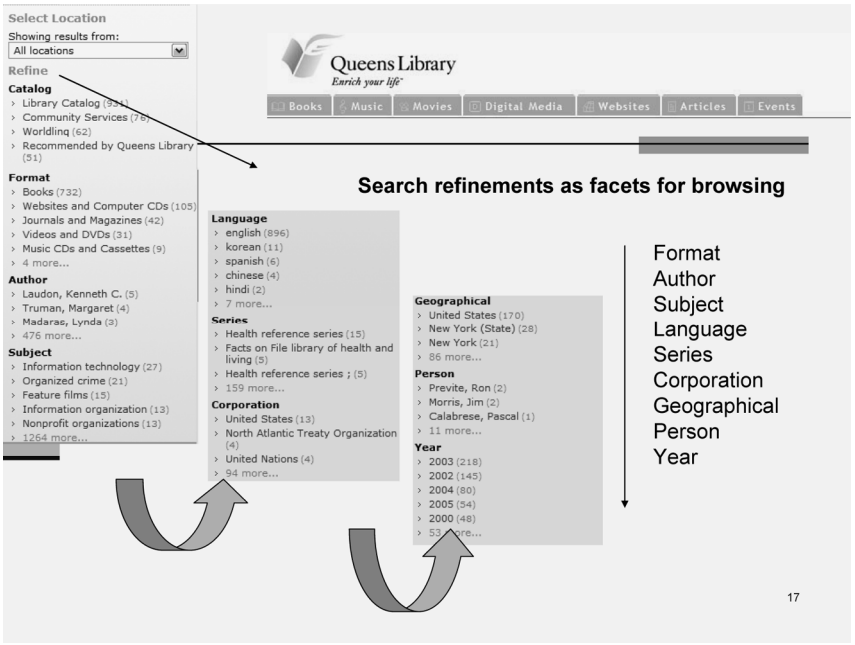

Figure 6. *Full display of facets provided by the Queens Library interface. (Screen capture, June 2007).*

Libraries will need the dynamic support offered by evaluative guidelines as they begin to assess, evaluate, design and implement new interfaces. Observing the effectiveness of these tools, and identifying critical system features that assist information seeking, or that are missing in these early implementations will better serve the widescale implementations that may soon follow. A research agenda that would address these needs and concerns is proposed here.

3. Research Agenda

Recent OPAC implementations have generated a groundswell of interest and excitement about re-

Application	Example	Purpose
Special subject schemes (Classification Research Group members)	*English Electric Scheme* (Binns and Bagley, 1958, 1961) *British Catalogue of Music* (1960) *Classification of Enterprise Activities* (1966) [Vickery, 1966]	From practice to theory (Richmond 1988)
Audacious (Atherton Cochrane and others)	Test of *UDC* as mechanized searching language (Atherton and Freeman 1967, Freeman and Atherton 1968a, 1968b, 1969)	Facets in computerized indexing and retrieval
View-based searching (Pollitt and others)	Hibrowse for Embase http://www.ifla.org/IV/ifla63/63polst.pdf (Pollitt, S. et al. 1994, 1996, 1998)	Facets support query formation and expansion by browsing
Flamenco (Hearst and others)	http://flamenco.berkeley.edu/tutorials.html (Hearst, M. et al. 2002, 2006a, 2006b)	Faceted metadata (automatic generation of facets)
FATKS (Broughton and others)	*Bliss classification* (2nd ed.) to provide access in a humanities subject portal http://www.ucl.ac.uk/fatks/ (Hockey Broughton and Slavic 2004)	Use of a faceted classification as a thesaurus and subject access device

Table 1. *Exemplars of traditional facet analysis*

placing current systems. Before the excitement turns into widescale implementations, some cautionary tales are warranted. The use of the term "facets" in the discussions about these new interfaces is strikingly similar to the term as used by information architects and knowledge managers just beginning to explore the design and implementation of faceted interfaces in 2000 (La Barre 2006). The information architects and knowledge managers interviewed as part of the preceding study expressed deep interest in use and user studies, but were unable to find the funding or the time needed in fast-paced commercial and business environments. In the case of use or user studies that were supported, results are often confidential, proprietary, for internal use, or are released anecdotally in the form of conference presentations, not as published research studies complete with findings and data. Design and implementation teams for library systems have been generous in discussing their experiences at recent conferences hosted by the American Library Association and other professional organizations and by publishing these in the professional literature (Antelman, Lynema, and Pace 2006; Collins, Samples, Pennell and Goldsmith in press). More comprehensive approaches that have been independently conducted are needed in order to assure comparability across user groups and implementations. Inquiry should be foregrounded by the traditional objectives of the catalog (Cutter 1876) especially as they are embodied in *Functional Requirements for Bibliographic Records* (*FRBR*) and *Functional Requirements for Authority Data* (*FRAD*) (IFLA 1997, 2007), special attention should be given to validating the *FRBR* user tasks. As *Resource Description and Access* (*RDA*) nears publication in 2009, this mission becomes even more critical.

As faceted search and discovery tools grow increasingly common on the Web, and it appears that they may become ubiquitous in library search and discovery environments, critical examination of the features of these implementations and the means by which they support or hinder searching is imperative. It appears that some of the systems currently being implemented in libraries are heavily influenced by the commercial and business applications that preceded them. In some cases libraries are using software first designed for use in e-commerce applications. Now is the time to question the assumptions that are embedded in these commercial systems that were primarily designed to provide access to concrete items through descriptions in order to enhance profit. While libraries provide access to items through descriptions of these items, their motivations are public service-oriented. Descriptions of library materials include analysis of the intellectual content, or concepts, embedded in the content of the items they make available. Data assigned to cultural objects, especially as they make use of name and subject authority data, are more complex and powerful than most metadata routinely assigned to commercial products.

In order to fully utilize the power of facets, facet analysis must be properly conducted prior to system implementation. Moreover, La Barre (2006) found that the majority of facets in use on a stratified random sample of 200 websites were superficially implemented. Several of the interviewees attributed this to the need to use pre-existing metadata or "low-hanging fruit" that could be easily harnessed for system use. Will these new systems allow these data to "work harder" without requiring changing the way in which they are created? Can facets and facet analysis, part of the legacy of LIS information access and discovery systems, be used to greater effect in OPACs? Are there other system features that would support searching that have not yet been considered, or remain undiscovered in the excitement of new implementations?

3.1 Study rationale

Many heterogeneous user groups utilize discovery tools in libraries, museums and archives in order to access cultural objects. Rather than focus on the "typical" undergraduate user, in keeping with the discussions emanating from the Library of Congress and elsewhere. It is critically important to focus on the scholarly user for several reasons. Some have suggested that choosing the typical user (an undergraduate student) is a short-sighted and illogical approach owing mainly to the fact that this captive and easily studied group has been over-studied. Others have suggested that perhaps libraries should abandon efforts to "bring back" the typical user and focus instead on the academic or scholarly user as a "niche group" that has always been well served by the library and has long ranked among the most steadfast users of library services (Mann 2006a, 2006b). Important lessons can be learned from this user group that can, in turn, improve services for all users.

3.2 Study goals and design

What is needed now is a three stage exploration into the information seeking behavior of scholarly researchers and their interactions with the new faceted browsing and navigation systems that are being implemented in order to support information access and discovery. The overarching research questions are derived from current debates about the best means by which to provide information seekers with robust information discovery and access tools. One goal of this study is to identify interface factors that hinder scholarly access to information. A second goal of this study is to identify features of information access and discovery tools that support scholarly research. A third goal of this study will be to categorize current practices of scholarly information seekers during their interactions with newly designed information access and discovery tools to determine whether there are features that would better meet their information seeking needs that are not currently available.

Participants could be selected from among faculty and students by using as selection criteria, their membership in one of the five disciplinary areas used by the JISC and RIN studies ((Sparks 2005, Research Information Network 2006): (1) medical and biological sciences, (2) physical sciences and engineering, (3) social sciences, (4) languages and area studies, and (5) arts and humanities. framework of disciplines. Identification of specific subject disciplines from which to draw participants will be partially dependent on the existence and availability of faceted search and discovery tools for each selected subject, though a list of these is currently being maintained.

By undertaking an interwoven approach to the study of information seeking and the information retrieval systems designed to support this process, an integrated conceptual framework will also be developed in order to scaffold more robust solutions to assist information seekers as suggested by Kuhlthau (2005) and others. Along the way assumptions built into models of information access and designs for discovery systems will be examined in order to develop suggestions for better ways to support users' discovery and access in libraries, museums and archives.

3.3 Study method

Interview and observation sessions will elicit several recent information seeking instances from each participant during the course of an interview structured using the critical incident technique (CIT). This open-ended method can assist in the determination of critical features for a system or interface by allowing for a direct focus on user behavior (Flanagan 1954, Carlisle 1986). Prior to the elicitation of incidents, the information seeking activity of interest will be carefully defined, and a structured set of interview prompts formulated in order to assure data consistency and validity throughout the data collection process. These prompts will also serve to assist

each participant to consistently recall information seeking incidents by asking them to (1) focus on an incident with a strong positive (or negative) influence on the information seeking incident, and describe it (2) to describe what led up to the incident. Responses will be recorded and transcripts of the interviews will be created.

In order to address the retrospective bias that is potential in CIT, each participant will also be presented a set of up to three new information discovery and access tools that are specific to the participant's discipline and which are representative of those provided by research libraries. Each participant will be asked to conduct a search using the information seeking incidents they have recalled. Through the use of *UserVue* web-based software (which captures moving images of the screen during system use instead of the participants themselves), website navigation processes can be viewed remotely and subject to video capture. This will allow observations to take place with the computer equipment and in a location that is familiar to each participant. Interviews will continue throughout the observation process by use of a think-aloud protocol in which each participant will be asked to discuss their information search as it unfolds. Wireframes will be created for each search and discovery tool that is presented to the subjects, and for comparative purposes, wireframes will be created for all search and discovery tools used by the subjects during their critical incidents. A coding manual will be created to assist in training wireframe and video capture coders, and intercoder reliability testing will be conducted after the first few transcripts, and periodically thereafter in order to ensure consistent, reliable and valid results.

Other studies of information seeking have relied heavily on interviews or survey methods to elicit information. It is anticipated that the use of CIT to structure interviews and the observations of participant interactions with faceted search and discovery tools, in combination with the critical incidents themselves and the descriptions of successes and frustrations with familiar search environments, will result in a rich data set for comparative analysis. A visit to the UK to meet with RIN and JISC researchers would also allow for discussion and comparison of study data. A comprehensive literature review of previous user or use studies that focus on scholarly information seeking will be the first step for the proposed study in order to inform the selection of subjects, search and discovery tools, coding manuals and interview structure. It will also assist later in the creation of taxonomy of observed information seeking. The data extracted from previous studies, and the data generated by the proposed study, will be interwoven in the construction of a conceptual framework which integrates the information seeking behavior of the scholars in the study, with the features and design of the systems being studied. This conceptual framework will assist in principled suggestions for feature improvements, and may also serve to confirm or disconfirm the benefits of these new faceted search and discovery systems for scholarly information seeking, assist in the evaluation and assessment of these systems, and potentially scaffold the creation of new system features, or improvements to existing features thereby leading to systemic change in the search and discovery systems available in libraries, archives and museums during this intensive period of system evaluation and implementation.

4. Concluding remarks

During this time of re-evaluation and redefinition about the future roles of legacy information access systems, fears are mounting: that legacy access systems may be a casualty of the brave new digital world, that librarians will become irrelevant, that users are becoming increasingly reliant on incomplete or irrelevant works easily found on the Internet while at the same time, missing quality information resources because of a mistaken assumption that using a search engine is a comprehensive search. Several recent articles serve as reminders that even though the principles guiding the creation of legacy library systems used technology that is now out of date, the principles that underpin them are more vital than ever (Mann 2007, Miller and Pellen 2006).

By providing support for feature analysis of search and discovery systems which views these systems from a scholarly information seeking perspective, evaluative guidelines can be constructed that may assist in the assessment of current systems and the development of new systems and comprehensive evaluative approaches for all users. By maintaining comparability with previous studies, and foregrounding traditional principles and objectives, it is hoped that a set of features that are useful to scholars and researchers may be suggested for implementation across the international library, archive and museum communities in time to enhance the utility and usability of the next generation of discovery and access environments.

Acknowledgments

Thanks for their comments on this work go to the referees and participants of the first North American Symposium for Knowledge Organization (NASKO). I would especially like to thank Rebecca Green, Mark Lindner, Claire Beghtol and Pauline Cochrane.

References

Anderson, James D. and Pérez-Carballo, José. 2005. *Information retrieval design: principles and options for information description, organization, display, and access in information retrieval databases, digital libraries, catalogs, and indexes.* East Brunswick, N.J.: Omecteca Institute.

Antelman, Kristin, Lynema, Emily, and Pace, Andrew K. 2006. Toward a 21st century library catalog. *Information technology and libraries* 25: 128-39.

Atherton, Pauline and Freeman, Robert R. 1967. *File organization and search strategy using the Universal Decimal Classification in mechanized reference retrieval systems* (AIP/UDC-5). New York: American Institute of Physics: UDC Project.

Bates, Marcia J., Wilde, Deborah N. and Siegfried, Susan. 1995. Research practices of humanities scholars in an online environment: The Getty online searching project report No. 3. *Library and information science research* 17(5): 40-58.

Bibliographic Services Task Force. 2005. *Rethinking how we provide bibliographic services for the University of California.* http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf (accessed 27 June 2007).

Binns, Jean. (undated, unpublished, manuscript). *A faceted classification scheme: construction and operation.* Held as part of the Cochrane papers at Syracuse University archives. (Notation indicates "Classification Research Study Group holdings").

Binns, Jean. 1961. *Operating experience with a facet system in the English Electric Company.* Unpublished manuscript held as part of the Cochrane papers at Syracuse University Archives.

Binns, Jean, and Bagley, D. 1958/1961. *A faceted classification for engineering.* 2nd/3rd ed. Whetstone: English Electric Company. Ltd.

Broughton, Vanda. 2001. Faceted classification as a basis for knowledge organisation in a digital environment: *The Bliss Bibliographic Classification* as a model for vocabulary management and the creation of multidimensional knowledge structures. *New review of hypermedia and multimedia* 7: 67-102.

Broughton, Vanda. 2002. Faceted classification as a basis for knowledge organization in a digital environment; *The Bliss Bibliographic Classification* as a model for vocabulary management and the creation of multidimensional knowledge structures. In López-Huertas, Maria-J. and Muñoz-Fernández, F. eds. *Challenges in knowledge representation and organization for the 21st century, Integration of knowledge across boundaries, Proceedings of the 7th International ISKO Conference, 10-13 July 2002, Granada, Spain*. Würzburg: Ergon, pp. 135-42.

Broughton, Vanda. 2004. Faceted classification: a tool for subject access in the twenty first century. *Signum* 8: 5-18.

Broughton, Vanda and Lane, Heather. 2000. Classification schemes revisited: applications to Web indexing and searching. *Journal of internet cataloging* 2(3/4): 143-55.

Byrd, Jackie et al. 2006. *A white paper on the future of cataloging at Indiana University.* http://www.iub.edu/~libtserv/pub/Future_of_Cataloging_White_Paper.pdf (accessed 27 June 2007).

Calhoun, Karen. 2006. The changing nature of the catalog and its integration with other discovery tools. Prepared for the Library of Congress. http://www.loc.gov/catdir/calhoun-report-final.pdf (accessed 27 June 2007).

Carlisle, Kenneth Eugene. 1986. *Analyzing jobs and tasks*. Techniques in training and performance development series. Englewood Cliffs, NJ: Educational Technology Publications, Inc.

Cochrane, Pauline. 1978. *Books are for use. Final report of the Subject Access Project to the Council on Library Resources.* ERIC Document 156131.

Cochrane, Pauline. 1981. Improving the quality of information retrieval – online to a library catalog or other access service … or … where do we go from here? *Online* 5(3): 30-41.

Collins, Maria, Samples, Jacquie, Pennell, Charley and Goldsmith, David. 2007. Magnifying the ILS with Endeca. *Serials librarian* 51(3/4): 75-100.

Cutter, Charles Ammi. 1876. *Rules for a printed dictionary catalogue*. Washington, D.C.: Government Printing Office.

Dempsey, Lorcan. 2005. Making data work harder. *Lorcan Dempsey's weblog on libraries, services and networks*. http://orweblog.oclc.org/archives/000535.html (accessed 27 June 2007).

Fast, Karl V. and Campbell, D. Grant. 2004. 'I still prefer Google': university student perceptions of

searching OPACs and the Web. In *Managing and enhancing information: cultures and conflicts, Proceedings of the 67th Annual Meeting of the American Society for Information Science and Technology, Providence, Rhode Island, November 13-18, 2004*. Medford, N.J.: Information Today, pp. 138-46.

Flanagan, John C. 1954. The critical incident technique. *Psychological bulletin* 51: 327-59.

Freeman, Robert R. and Atherton, Pauline. 1968a. AUDACIOUS: An experiment with online, interactive reference retrieval system using the *Universal Decimal Classification* as the index language in the field of nuclear Science. AIP.UDC-7. New York: American Institute of Physics, UDC Project.

Freeman, Robert R. and Atherton, P. 1968b. Final report of the research project for the evaluation of the *UDC* as the indexing language for a mechanized reference retrieval system. Springfield, VA: American Institute of Physics.

Freeman, Robert R. and Atherton, Pauline. 1969. Proceedings of a first seminar the UDC in a mechanized retrieval system. In: R. Moelgaard-Hansen and Rigby, Malcolm eds, *Committee on Classification Research Seminar*. FID/CR Report No. 9. Serie-No. 405. Copenhagen: Danish Center for Documentation.

Friedlander, Amy. 2002. *Dimensions and use of the scholarly information environment: introduction to a data set.* DLF/CLIR pub110. *http://www.clir.org/pubs/abstract/pub110abst.html* (accessed 27 June 2007).

Hearst, Marti A. 2006. Clustering versus faceted categories for information exploration. *Communications of the ACM* 49(4): 59-61.

Hearst, Marti A., Elliott, Ame, English, Jennifer, Sinha, Rashimi, Swearingen, Kirsten, and Lee, Ka-Ping. 2002. Finding the flow in web site search: designing a search system and interface may best be served (and executed) by scrutinizing usability studies. *Communications of the ACM*, 45(9): 42-49.

Hearst, Marti A., Smalley, Preston and Chandler, Corey. 2006. *Faceted metadata for information architecture and search.* CHI 2006 course. http://bailando.sims.berkeley.edu/talks/chi_course06.pdf (accessed 27 June 2007).

Hildreth, Charles R. 1982. *Online public access catalogs: the user interface.* Dublin, OH: OCLC Online Computer Library Center.

Hildreth, Charles R. 1995. *Online catalog design models: are we moving in the right direction?* Washington, D.C.: Commissioned by the Council on Library Resources. http://myweb.cwpost.liu.edu/childret/clr-opac.html (accessed 27 June 2007).

Hockey, Susan, Broughton, Vanda and Slavic, Aida. 2004. Facet analytical theory in managing knowledge structures for the humanities. (FATKS). http://www.ucl.ac.uk/fatks/fat.htm (accessed 27 June 2007).

Horrigan, John B. and Rainie, Lee. 2002. *Counting on the Internet.* Washington, D.C.: Pew Internet & American Life Project. http://www.pewinternet.org/reports/toc.asp?Report=80 (accessed 27 June 2007).

IFLA Study Group on the Functional Requirements for Bibliographic Records. 1998. *Functional requirements for bibliographic records: final report*. München: K.G. Saur. http://www.ifla.org/VII/s13/frbr/frbr.pdf (accessed 27 June 2007).

IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). 2007. Functional requirements for authority data: a conceptual model. Draft 2007-04-01. http://www.ifla.org/VII/d4/FRANAR-ConceptualModel-2ndReview.pdf (accessed 27 June 2007).

Joint Information Systems Committee (JISC). 2006. http://www.jisc.ac.uk (accessed 27 June 2007).

Kuhlthau, Carol C. 2005. Towards collaboration between information seeking and information retrieval. *Information research* 10(2). http://informationr.net/ir/10-2/paper225.html (accessed 27 June 2007).

La Barre, Kathryn. 2006. The use of faceted analytico-synthetic theory as revealed in the practice of website construction and design. Ph.D. dissertation, Indiana University.

Mann, Thomas. 2006a. A critical review [of] The changing nature of the catalog and its integration with other discovery tools. Final Report." March 17, 2006. Prepared for the Library of Congress by Karen Calhoun. http://www.guild2910.org/AFSCMECalhounReviewREV.pdf (accessed 27 June 2007).

Mann, Thomas. 2006b. What is going on at the Library of Congress? http://www.guild2910.org/AFSCMEWhatIsGoingOn.pdf (accessed 27 June 2007).

Mann, Thomas. 2007. The Peloponnesian war and the future of reference, cataloging, and scholarship in research libraries. http://guild2910.org/Pelopponesian%20War%20June%2013%202007.pdf (accessed 27 June 2007).

Marcum, Deanna B. 2005. The future of cataloging: address to the Ebsco Leadership Seminar, Boston,

Massachusetts, January 16, 2005. http://www.loc.gov/library/reports/CatalogingSpeech.pdf (accessed 27 June 2007).

Meho, Lokman I. and Tibbo, Helen R. 2003. Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology* 54: 570-87.

Miller, William and Pellen, Rita M. eds. 2006. *Libraries and Google*. Binghampton, N.Y.: Haworth Press.

Murphy, Janet. 2003. Information-seeking habits of environmental scientists: A study of interdisciplinary scientists at the environmental protection agency in Research Triangle Park, North Carolina. *Issues in science & technology librarianship* 38. http://www.istl.org/03-summer/refereed.html (accessed 12 July 2007).

North Carolina State University. 2006. *Endeca at the NCSU librar*ies. http://www.lib.ncsu.edu/endeca/ (accessed 27 June 2007).

Nelsonville Public Library. https://www.athenscounty.lib.oh.us/ (accessed 27 June 2007).

OCLC Online Computer Library Center. 2002. *OCLC white paper on the information habits of college students*. Dublin, Ohio: OCLC Online Computer Library Center, Inc. http://www5.oclc.org/downloads/community/informationhabits.pdf (accessed 27 June 2007).

Palmer, Carole L. 1996. Information work at the boundaries of science: linking library services to research practices: findings of a study of scientists at an interdisciplinary institute. *Library trends* 45(2): 165-91.

Palmer, Carole L. and Neumann, Laura. 2002. The information work of interdisciplinary humanities scholars: exploration and translation. *Library quarterly* 72: 85-117.

Pollitt, A. Steven. 1998. The essential role of dynamic faceted classification in improving access to online information. *Proceedings of the 22nd International Online Information Meeting*. New York: Learned Information, pp. 17-21.

Pollitt, A. Steven., Ellis, Geoffrey P., and Smith, Martin P. 1994. HIBROWSE for Bibliographic Databases. *Journal of information science* 20: 413-26.

Pollitt, A. Steven, Smith, Martin P., Braekevelt, Patrick A.J. and Treglown, Mark. 1996. HIGHBROWSE for EMBASE: The application of view-based searching techniques to Europe's most important biomedical database. In *INFO 96 11th Annual Information Meeting, Tel Aviv, Israel May 5-9 1996*. Berlin: Springer-Verlag.

Queens Library Aquabrowser Library. http://aqua.queenslibrary.org/ (accessed 27 June 2007).

Ranganathan, S.R. 1962. *Elements of library classification, based on lectures delivered at the University of Bombay in December 1944 and in the schools of librarianship in Great Britain in December 1956*. Bombay: Asia Publishing House.

Richmond, Phyllis A. 1988. Precedent-setting contributions to modern classification. *Journal of documentation* 44: 242-49.

Research Information Network. 2006. Researchers and discovery services: Behaviour, perceptions and needs. http://www.rin.ac.uk/files/Report%20-%20final.pdf (accessed 27 June 2007).

Research Information Network (RIN). http://www.rin.ac.uk/ (accessed 27 June 2007).

Saracevic, Tefko. 1997. Users lost: reflections on the past, future, and limits of information science. SIGIR Forum 31(2): 16-27.

Sparks, Sue. 2005. JISC disciplinary differences report. London: Rightscom. http://www.jisc.ac.uk/uploaded_documents/Disciplinary%20Differences%20and%20Needs.doc (accessed 27 June 2007).

Tenopir, Carol. 2003. User and users of electronic *library resources: an overview and analysis of recent research studies.* With the assistance of Brenda Hitchcock and Ashley Pillow. CLIR pub. 120. Washington, D.C.: Council on Library and Information Resources. http://www.clir.org/PUBS/reports/pub120/pub120.pdf (accessed 27 June 2007).

Unsworth, John. 2000. Scholarly primitives: what methods do humanities researchers have in common, and how might our tools reflect this?" In *Symposium on Humanities Computing: Formal Methods, Experimental Practice*, King's College, London, 13 May 2000. http://www.iath.virginia.edu/~jmu2m/Kings.5-00/primitives.html (accessed 27 June 2007).

Vickery, B. C. 1966. Faceted classification schemes. Rutgers series on systems for the intellectual organization of information v. 5. New Brunswick, NJ: Graduate School of Library Service, Rutgers, the State University.

Wilson, T.D. 2000. Human information behavior. *Informing science, 3*(2): 49-55. http://inform.nu/Articles/Vol3/v3n2p49-56.pdf (accessed 27 June 2007).