The Metadata *is* the Interface
Better Description for Better Discovery of Archives and Special Collections, Synthesized from User Studies

Jennifer Schaffner

Program Officer
OCLC Research

OCLC

Suggested citation:
Schaffner, Jennifer. 2009. "The Metadata is the Interface: Better Description for Better Discovery of Archives and Special Collections, Synthesized from User Studies." Dublin, OH: OCLC Research. https://doi.org/10.25333/dp1k-3348.

Contents

Introduction

Tim Ericson warned that user studies are important, but "they can also be a substitute for more direct action."[1] We have strong evidence about how to improve discovery of archives and special collections, and we need to start somewhere. These days we are writing finding aids and cataloging collections largely to be discovered by search engines. People expect to find archives and special collections on the open Web using the same techniques they use to find other things, and they expect comprehensive results. Invisibility of archives, manuscripts and special collections may well have more to do with the metadata we create than with the interfaces we build. Now that we no longer control discovery, the metadata that we contribute is critical. In so many ways, the metadata *is* the interface.[2]

Structured metadata can be useful internally for collection management and public services, but is not always what users need most to discover primary sources, especially minimally-described collections and "hidden collections."[3] We understand archival standards for description and cataloging, but our users by and large don't.[4] Studies show that users often do not want to search for collections by provenance, for example, as important as this principle is for archival collections.[5] One of several core competencies that special collections metadata librarians must have is "a keen understanding of users' needs and preferences."[6] This is especially important now that discovery happens in multiple environments.[7] Librarians and archivists need to *manage* archival collections by provenance, but also must describe what is *in* the collections for their users.

This essay—part of a series of OCLC Research projects to mobilize unique materials—synthesizes evidence of what descriptive information people say they need for research.[8] As this literature review got underway, it soon became evident that we already know most of what we need to know in order to get started making changes.

In many contexts over many years, librarians and archivists have studied users with a wide variety of research methods: using surveys and questionnaires, examining statistics and citations, testing usability of interfaces, studying information-seeking behaviors, listening to focus groups, creating personas, and questioning the efficacy of finding aid portals.[9] The goal has always been to improve practices in order to help people—not just archivists and librarians—discover archival and rare materials.[10] We still have gaps in our understanding, and comparing different kinds of studies across many years of work is like comparing apples and oranges. Nevertheless, the community has learned from these studies about obstacles between people and unique materials. While there is

more to learn, let's start *now* by adjusting our practices in order to disclose information about special collections and archives more effectively.

Librarians and Archivists as Gatekeepers

Users work increasingly on their own, while librarians and archivists have expected to mediate research. Most often people want to be autonomous and discover information about primary sources at the network level, not at the institutional level.[11] In an Ithaka study of higher education, Roger Schoenfeld and Ross Housewright learned that scholars consider *less* mediation in research and discovery a good thing:

> [L]eading-edge libraries are beginning to change their priorities to match those of faculty and students. Still, the mismatch in views on the gateway function is a cause for further reflection: if librarians view this function as critical, but faculty in certain disciplines find it to be declining in importance, how can libraries, individually or collectively, strategically realign the services that support the gateway function?"[12]

The more that discovery occurs directly via search engines, the greater the success of considerable efforts to expose "hidden collections."

Over twenty-five years ago, Mary Jo Pugh challenged the myth of immortal and omniscient archivists, on whom users would rely for access to the contents of archival collections.[13] Many studies of library catalogs and archival portals have shown that these days most users start their search for information with Google or Wikipedia, and usually only come to libraries and archives for known items.[14] Now the primary role in discovery is making the collections more visible and staying out of the way:

> "Perceptions of a decline in dependence are probably unavoidable as services are increasingly provided remotely, and in some ways these shifting faculty attitudes can be viewed as a sign of library success. One can argue that the library is serving faculty well, providing them with a less mediated research workflow and greater ability to perform their work more quickly and effectively. In the process, however, they may be making their own role less visible."[15]

Perhaps ironically, goals to disclose descriptions online and to digitize primary resources have made special collections more visible and roles of archivists and librarians less visible. The more users do not need to consult archivists and librarians for searching, the more successful initiatives to improve description and discovery have been.

Users Search by Subjects and Keywords

Archivists and librarians have often focused on what collections are made up of (*Ofness*), while many users prefer to learn what collections are about (*Aboutness*).[16] Studies report consistently that many users want to find information about contents of collections.[17] For instance, Bill Maher analyzed reference letters to the University of Illinois archives in 1984-85 and found that over one third of the researchers inquired about subjects.[18] One respondent in Jane Stevenson's testing of the UK's Archives Hub said, "I like the subject finder. I'm pleasantly surprised by it."[19] In the most recent Northwest Digital Archives (NWDA) usability test, one user was enthusiastic to discover the subject section: "These will give me an idea of what this collection is about."[20] In a previous NWDA usability study, one person recommended controlled subject vocabulary and wanted subject terms linked to other collections and catalogs.[21] Wendy Duff concluded, in more than one study, that users wanted "what is it about?" to appear at first glance.[22] Louise Gagnon-Arguin found 41% of queries in Québecois archives were for subjects or themes.[23] A study in 1976 of registration forms at the Michigan Historical Society showed that, "Roughly half of all users, regardless of preparation, began with a subject searching approach."[24] For thirty years, people have reported that they want to discover archival materials using subject information.

Content is more important than format.[25] Over fifteen years ago, Jackie Dooley cautioned that without subject access to records about archival collections, users are reduced to known-item searching.[26] An example of this surfaced in recent usability testing of WorldCat Local at the University of California (UC). Faculty and graduate student participants only searched UC's union catalog for known items, not for discovery, when they were working in their areas of expertise.[27] In an example of good intentions, the Online Archive of California (OAC) hoped to add subject searching until they learned that only "60% of the finding aids used controlled access tags."[28] Richard Szary and Lawrence Dowler recommended "direct indexing of the content of historical materials" to improve access.[29] For discovery, *Aboutness* is a very important element of description.

While users want to find subjects, they generally search using keyword techniques, rather than by using structured terminology. For example, research shows that keywords are important to historians searching for known items.[30] Likewise, NWDA usability testers observed that searches were completely unstructured.[31] In November 2008, the French CALAMES project reported 40% frequency of searching full text, 34% by personal name, and 19% by various subject elements and attributes.[32] Susan Hamburger's research yielded different proportions: 78% by keywords, 31% by names and 23% by subjects.[33] Chris Prom also found that users of the University of Illinois's electronic finding aids primarily used non-fielded keyword search terms, along with structured browsing.[34]

Recent work addressed phrase-searching techniques. Phrase searches have been shown to be more effective than keyword searches when using search engines to find finding aids.[35] People don't search that way, however, according many studies, including OAC usability testing in 2001.[36] In another example, 8 out of 9 participants searched by keywords—not phrases—in NWDA usability testing.[37] Kristina Southwell used statistical reports from search engines to demonstrate that the University of Oklahoma's Web pages for manuscripts were typically found through keyword searches, although some people used subject phrases, too.[38] Based on research with users at six major research libraries, Susan Hamburger recommended offering searching on both keywords and subject terms in catalog records and finding aids.[39]

A wide range of research shows that keyword searching is important specifically for humanities scholars, who often search using name, place, title and discipline-specific terms.[40] Jihyun Kim examined EAD finding aids themselves, rather than users, precisely because historians and humanists search for primary sources by names of people and places. Kim reported that few finding aids used "controlled access headings."[41] Wendy Duff and Catherine Johnson interviewed ten historians and concluded they search names primarily because names are the easiest way into collections. Social historians desired subject access to collections: "'There has to be a way that people can find things without having to know who generates them, so keywords will search across different provenances of things' (participant 6)."[42] Using keyword searching techniques for topics—such as farm women—can be problematic, because archives are organized primarily by the names of the creators, not the subject content of the collections.

There is no common understanding of what users and testers mean when they use words like "keyword," "subject," "known item," "name," "phrase" and "browse." Without that common understanding, it is difficult to compare findings from separate studies. Is a keyword search technique in effect a subject search, from a user's standpoint?[43] In one test, while Archives Hub participants favored subject searching, they were confused by a browse list composed of access points.[44] Do testers consider natural language searching to be keyword or subject searching, even if the user's search includes names? Wendy Duff and Catherine Johnson, for example, consider a search by name keywords to be a known-item search.[45] Users do not always distinguish clearly between names and subjects. For instance, two of the participants in the Archives of American Art usability study never found the Joseph Cornell collection because they searched by keywords rather than browsing an alphabetical list of collections.[46] RLG learned from focus groups that many participants combine keywords with names, subjects and dates.[47] Richard Lytle speculated that many kinds of searches might be disguised subject searches: "Requests for records by proper name, geographical area, date or form may conceal a subject request. Does the user really prefer to ask for documents by name…?"[48]

Elsie Freeman memorably posited that good subject information is a large component of discovery experiences that are simple, elegant and intuitive.[49] Users want to search names by keyword, search for subjects by browsing, and browse by keyword or name, too. When it comes to using descriptive metadata to discover archival materials and special collections, users want it all. This is problematic because significant principles of archival theory and practice have been provenance and description of what the collection is made up of, its *Ofness*.[50] For users, research shows that important elements of description, especially minimum-level description, are keywords and terms that indicate *Aboutness*.

Users Expect Results Ranked by Relevance

While researchers consider it important to know the relative importance of collections, archivists and librarians rarely create metadata that can be used to rank relevance. In 1987, Avra Michelson argued that scholars using primary sources expected relevant results when doing research in exhaustive listings of collections.[51] Over twenty years later, students at the University of Maryland were overwhelmed by large result sets retrieved by keyword searches; they expected relevance ranking of results such as that returned by Google and other search engines.[52] Chris Prom learned—using the interface for the University of Illinois Archives—that hits sorted by provenance confused his participants, who were largely expecting search results to be ranked by relevance.[53] Andrea Rosenbusch concluded, after studying a dozen archival online databases, that, "The relevance of provenance as the main access point to records is becoming questionable…"[54]

As it stands now, identifying relevant primary resources often requires educated guesswork. All of the participants in Sara Snyder's study at the Archives of American Art said that relevance ranking was essential, especially for large results.[55] On the other hand, in Jane Stevenson's Archives Hub study, relevance ranking of the results of a subject search puzzled some people, who then wanted to know how relevance worked and why some hits were more relevant.[56] When redesigning ArchiveGrid for improved usability, RLG determined that the order of search results was important, and relevance—not title—was the desired order.[57]

Several tactics have been proposed that could indicate the relative importance of special collections in discovery experiences. Extent or physical description elements can be useful for some researchers trying to sort out relevance for themselves. ("Just one quick question. Does anybody understand what twelve metres of textual records means?" "Means two weeks in the archives!"[58]) Andrea Rosenbusch suggested relevance could be leveraged from multi-level description, by restricting queries to top-level descriptions: "The aim [of ISAD(G)] is to enable users to identify *fonds* or even whole collections which have the highest relevancy to them."[59] Systems don't exist yet that use standards-based descriptions and extent statements in this way.

Search engine optimization strategies could leverage metadata for sorting search results by relevance. Based on keyword density analysis of UC Irvine's finding aids, Michelle Light advocated enhancing discovery by describing collections more strategically—by using more keywords and concepts than folder lists and material types.[60] Taking another tactic, the NWDA Working Group recommended experimenting with algorithms to combine use statistics with the frequency of index terms in order to produce relevance ranking like in search engines.[61] Recommender systems for discovery of archival collections might provide indications of relevance. Improvements will require imaginative use of available Web 2.0 tools, such as tags for important collections on a topic, or "link paths" like those demonstrated in the Polar Bear project.[62]

Over twenty years ago, Avra Michelson called for study of search questions, in order to identify successful patterns. Michelson recommended subsequent improvements in our use of subject terms in description in order to improve what she called "retrieval capabilities."[63] More recently, Karen Markey has similarly suggested we would learn a great deal from studying people's search terms.[64] Many user studies for archives and special collections have focused on discovery within local systems designed for archival materials.[65] Now that close to 90% of searching behavior begins in search engines,[66] it is time to evaluate search behaviors at the network level, in order to develop descriptive strategies for ranking the relevance of primary resources.

Comprehensive Coverage

Increasingly, archivists and librarians are acutely aware that many researchers expect comprehensive coverage. A student in the Maryland study expected that "the universe of primary sources is a finite, absolute body of material that can and has been already labeled and categorized for him."[67] Chris Prom, too, learned that many inexperienced users assume that everything is available.[68] Jane Stevenson's study with Archives Hub confirmed that some people assumed their search results were comprehensive.[69] In a usability study of the Lilly Library's Web site, Erika Dowell found that users doubted the utility of the online catalog when cautioned (responsibly) that only 45% of the Lilly's holdings were included.[70] In a related study in UK museums, the Research Information Network concluded that "what researchers need above all is online access to the records in museum and collection databases to be provided as quickly as possible, whatever the perceived imperfections or gaps in the records."[71]

Some researchers have substantiated a "More Product, Less Processing" (MPLP) approach to description and digitization.[72] At the University of Wisconsin, Joshua Ranger and Krystyna Matusiak are experimenting with a less expensive, streamlined process for mass digitization of archival collections. The students they interviewed all preferred more description, not less. However, when the comparative costs of full and minimal records were explained, all of the participants said

streamlined description was preferred: "Better than not having it at all."[73] The American Heritage Center at the University of Wyoming surveyed 600 respondents for their satisfaction with minimal processing. Asked to rank archival priorities, respondents most often chose "putting more resources into creating basic descriptions for all collections."[74] The MPLP approach matches users' acceptance of minimum-level description because they would like to discover more materials online. When such decisions are made to describe more collections at a minimal level, archivists and librarians need to indicate, however briefly, what the contents of collections are about.

Users Know How to Scan and Scroll

Archivists and librarians have worried about confusing users by presenting different kinds and amounts of metadata, while users mostly care more about what is in the collections.[75] Diverging desires for less or more information appear often in user research. One example of this variety is that Maryland students were able to use long finding aids easily, despite difficulties navigating specific tools to discover them. On the other hand, in the same test one Maryland student reported that "too much information hindered the reading of the display."[76] In another contradictory example, some of Wendy Duff's participants preferred to see shorter abstracts and scope-and-content notes, disparaging long biographical notes (ranked 16th in order of preference) or administrative histories (ranked 23rd). However, a different participant in the same study said anyone interested in long notes can scroll down through the display.[77]

There are many more examples of preferences for both brief and for full displays that support arguments for both minimal and full description. In the RLG rapid iterative interface testing, most participants found a brief scope-and-content note most useful.[78] On the other hand, studies also report that users know how to skim long pages of records, when they want to. In Jane Lee's usability testing for the 2008 OAC redesign, she noted that they chose a long display format for search results because, as one participant said, "it's nice to have a little more information" when browsing.[79] In NWDA usability testing, "the majority of the users started the search for information by skimming or scrolling through the finding aid page; most said they weren't reading for content, rather were scanning for key terms."[80] Genealogists in RLG's Archival Resources focus groups preferred to scroll through large result sets.[81] These conflicting recommendations suggest that minimum description may come as a relief to some users, but others prefer a full description. If a collection is fortunate enough to have full description, it will not necessarily get in a user's way when scanning and scrolling through results. Users support concise minimum-level description, which can also be effective for discovery when it is done well.

Users' Lack of Awareness

Archivists and librarians have created catalogs and portals, but many users don't use them or don't know they exist. "The greatest barrier to use is lack of awareness."[82] Often it isn't easy to find rare and unique library and archival materials because successful discovery currently requires people to understand what they are looking for and how to find it.[83] Karen Markey says rare and unique materials are invisible: "Thousands of special collections that make up the invisible Web feature their own unique search engines because their content is not accessible via general Web search engines."[84] Louise Gagnon-Arguin concluded that the key to access is fragile in the context of electronic information.[85] In order to find primary resources, people need to know too much about how collections are described and where those descriptions are lodged. That isn't good enough.

Catalogs don't seem to do the trick. "It is unlikely that researchers approach doing research by looking for a tool for doing research."[86] In the RIN user study of UK museums, "most researchers are unaware of the online catalogues…"[87] Beth Yakel, Susan Hamburger, Bill Maher and others have found that the majority of researchers do not use utilities such as ArchivesUSA, OCLC, RLIN or NUCMC.[88] While a percentage of people in Kristina Southwell's Oklahoma survey found manuscript collections by searching the Web, only one person (0.4% in 230 responses!) used RLIN's AMC.[89] Southwell was surprised that only 11.3% of respondents discovered manuscript collections using the online catalog, leading her to wonder about the considerable investment creating MARC records. 17.9% found collections from html finding aids on the Web site, 25.1% used footnotes and bibliographies, while another 8.6% used a published guide to the repository for know-item discovery. Users may search on the open Web, but often they find archives indirectly.

So are finding aids best for discovery? Bill Maher questioned out loud our tacit belief that better finding aids will automatically result in better access.[90] Most participants in Jane Stevenson's study of the Archives Hub "did not mention any kind of cross-searching networks."[91] Kathleen Feeney concluded that "electronic finding aids may not be well suited to serve as pointers to archival collections," based on her 1999 study of retrieval of full-text finding aids by search engines. Feeney concluded that "MARC records remain a more valuable and reliable means of locating archival resources" because of problems with relevance rankings at the network level.[92]

For successful discovery, what are the lessons learned about our choices for description? Early on, Rob Spindler and Richard Pearce-Moses argued for adapting description methods—based on their case study with Arizona State University patrons—expressly to improve comprehension of AMC records in an integrated online environment.[93] More recently, Michelle Light suggested strategies to adapt description than can "enhance retrieval possibilities" at the network level: use long-tail keywords, repeat names and keywords (bending rules for description), put the most important

content at the top, say more with less.[94] If students now don't look in library catalogs or archival portals for primary materials, why spend resources that way? Let's put the right descriptive metadata in the right places.

Conclusion

I argue that some thirty years of user studies teach that *Aboutness* and relevance matter most for discovery of special collections, especially now that discovery happens elsewhere.[95] Unfortunately, there is a gap between the expectations of users and historical descriptive practices in archives and special collections. Changes must be made to description because researchers rarely look in library catalogs or archival portals for primary resources. These changes are even more important for collections that have been selected for minimal processing and description. Ensuring that "hidden collections" can be discovered requires appropriate description, not just expert processing, cataloging and cross-searching networks. It would be heartbreaking if special collections and archives remained invisible because they might not have the kinds of metadata that can easily be discovered by users on the open Web.

In a 1986 article on "The Use of User Studies," Bill Maher described archivists with instincts about how their collections are used—but without data to support their instincts—as "working in the dark."[96] Since then, research demonstrates recurring observations of users' needs and preferences when they search for special collections and archives. Over time, users have adapted their research tactics: from discovery only by visiting repositories and by consulting printed catalogs or guides, then discovery using online catalogs and portals, and now discovery on the Web. All along, user studies have demonstrated that descriptive metadata indicating *Aboutness* and relevance matters significantly for discovery. Twenty years later, we are not working in the dark any more.

Notes

1   Timothy L. Ericson, "Preoccupied with Our Own Gardens: Outreach and Archives," *Archivaria* 31 (winter 1990-1991): 120. Richard Cox expressed similar concerns: "User studies and citation analyses float isolated from any practical applications." See "Researching Archival Reference as an Information Function: Observations on Needs and Opportunities," *RQ* 31, no.3 (spring 1992): 393.

2   This aphorism was coined by Arnold Arcolio.

3   Judith M. Panitch conducted the first survey that catalyzed the community around the significance of "hidden collections": *Special Collections in ARL Libraries* (Washington DC: Association of Research Libraries, 2001). Published online at: http://www.arl.org/bm~doc/spec_colls_in_arl.pdf. See also Barbara M. Jones, "Hidden Collections, Scholarly Barriers" (June 6, 2003). Published online at: http://www.arl.org/bm~doc/hiddencollswhitepaperjun6.pdf.

4   I have borrowed this idea from Steve Hensen.

5   Wendy M. Duff and Catherine A. Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 477, footnote 5.

6   ACRL Rare Books and Manuscripts Section Task Force on Core Competencies for Special Collections Professionals, "RBMS Competencies for Special Collections Professionals," *College and Research Libraries News,* 69, no. 10 (November 2008): 622.

7   Lorcan Dempsey, "The Library Catalogue in the New Discovery Environment: Some Thoughts," *Ariadne* 48 (July 2006). Published online at: http://www.ariadne.ac.uk/issue48/dempsey/.

8   I focus attention on metadata for discovery, not on research methodologies, nor on interfaces, nor on interface terminology, nor on information-seeking behavior writ large, nor on what users do with what they find, much as these topics are fascinating and closely related to the relationship between description and discovery.

9   Many more studies have been done than are reported: "In recent years archive services in many parts of the world have undertaken surveys of their users. Valuable work has been done, but most of it suffers from lack of availability–many surveys remain unpublished and largely inaccessible…" Geoffrey Yeo, "Understanding Users and Use: A Market Segmentation Approach," *Journal of the Society of Archivists* 26, no. 1 (April 2005): 43-44.

10 Lawrence Dowler called for a national study of the use and users of archives to inform archival practices: "The Role of Use in Defining Archival Practice and Principles," *American Archivist* 51 (winter/spring 1988): 74-95. A stellar example is the National Council on Archives in the UK, which established standards requiring repositories to give users a choice in access to content, and to open up that content to a wider range of users: David Mander, ed., *Standard for Access to Archives* (National Council on Archives [UK]: Public Services Quality Group, 1st ed. 1999, 2nd ed. 2003, 3rd ed. 2008), 37, 46. The Auditor General of Canada is taking a similar tactic: Jean Dryden, "Do We Care What Users Want? Evaluating User Satisfaction and the LibQUAL+™ Experience," *Journal of Archival Organization* 4, no. 4 (2004): 83-84.

11 Louise Gagnon-Arguin applied the term "autonomy" to users doing research: "Les questions de recherché comme matériau d'études des usagers en vue du traitement des archives," *Archivaria* 46 (winter 1998): 86-102.

12 Ross Housewright and Roger Schoenfeld, *Ithaka's 2006 Studies of Key Stakeholders in the Digital Transformation in Higher Education* (August 2008). Published online at: http://www.ithaka.org/research/Ithakas%202006%20Studies%20of%20Key%20Stakeholders%20in%20the%20Digital%20Transformation%20in%20Higher%20Education.pdf.

13 Mary Jo Pugh, "The Illusion of Omniscience: Subject Access and the Reference Archivist," *American Archivist* 45, no. 1 (winter 1982): 38. Dowler also targeted mediation with humor: "To archivists, mediation has generally meant the satisfying vision of the erudite archivist leading a grateful scholar by the hand through

the uncharted forest of records to precisely the right material." "The Role of Use in Defining Archival Practice and Principles," *American Archivist* 51 (winter/spring 1988): 82.

[14] For example, Jennifer Ward, Pam Mojfeld, and Steve Shadle, "WorldCat Local at the University of Washington Libraries," *Library Technology Reports* 44, no. 6 (August /September 2008). Published online at: http://www.techsource.ala.org/ltr/worldcat-local-at-the-univ-of-washington-libraries.html.

[15] Ross Housewright and Roger Schoenfeld, *Ithaka's 2006 Studies of Key Stakeholders in the Digital Transformation in Higher Education* (August 2008). Published online at: http://www.ithaka.org/research/Ithakas%202006%20Studies%20of%20Key%20Stakeholders%20in%20the%20Digital%20Transformation%20in%20Higher%20Education.pdf.

[16] Sarah Shatforth (Layne) first applied the concepts *Ofness* and *Aboutness* to cataloging: "Analyzing the Subject of a Picture: a Theoretical Approach," *Cataloging & Classification Quarterly* 6 (1986): 39-62. Users and user testers have not used the terms.

[17] For example, a recent study of users of online library catalogs found that users desired "more subject information" second only to more online content: Karen Calhoun, *Online Catalogs: What Users and Librarians Want* (Dublin, Ohio: OCLC, 2009), 13, 43, 47. Published online at: http://www.oclc.org/reports/onlinecatalogs/fullreport.pdf.

[18] William J. Maher, "Use of User Studies," *Midwestern Archivist* 11 (1986): 21.

[19] Jane Stevenson, "'What Happens If I Click on This?' Experiences of the Archives Hub," *Ariadne* 57 (October 2008). Published online at: http://www.ariadne.ac.uk/issue57/stevenson/.

[20] Northwest Digital Archives (NWDA), "Executive Summary: Usability Test #5" (November 6, 2008), 6. Published online at: http://www.orbiscascade.org/index/usability-design-working-group-reports.

[21] NWDA, "Executive Summary: Usability Testing Round 4" (March 12, 2008), 6, 7.

[22] Wendy Duff, and Penka Stoyanova, "Transforming the Crazy Quilt: Archival Displays from the Users Point of View," *Archivaria* 45 (spring 1998): 63. "Many of the historians interviewed, particularly the social historians, suggested that subject indexes, keyword searches, or identification of themes would help them with their research": Wendy Duff and Catherine Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 490.

[23] 21% were by name, and 33% were by the type of document. Louise Gagnon-Arguin, "Les questions de recherché comme matériau d'études des usagers en vue du traitement des archives," *Archivaria* 46 (winter 1998): 92.

[24] Mary Jo Pugh, "The Illusion of Omniscience: Subject Access and the Reference Archivist," *American Archivist* 45, no. 1 (winter 1982): 40.

[25] Dana Chisnell, *Report: RLG Archival Resources Service Redesign Focus Groups* (Usability Works, 2 December 2004), 18. For similar results from recent user studies in a library context, see Karen Calhoun et al., *Online Catalogs: What Users and Librarians Want,* (Dublin, Ohio: OCLC, 2009). Published online at: http://www.oclc.org/reports/onlinecatalogs/fullreport.pdf. See also *2004 Information Format Trends: Content, Not Containers* (Dublin, Ohio: OCLC, 2004). Published online at: http://www5.oclc.org/downloads/community/2004infotrends_content.pdf.

[26] Jackie M. Dooley, "Subject Indexing in Context," *American Archivist* 55 (spring 1992): 351. Richard Smiraglia proposed methodology for subject access in "Subject Access to Archival Materials Using LCSH," *Cataloging & Classification Quarterly* 11, no. 3/4 (1990): 63-90.

[27] Arnold Arcolio and Felicia Poe, *WorldCat Local at the University of California: Usability Testing, Round One* (Spring 2008), 4.

[28] Michelle Light, "The Endangerment of Trees," EAD @ 10 (August 31, 2008): 1, *forthcoming.* Published online at: http://www.archivists.org/publications/proceedings/EAD@10/Light-EAD@10.pdf.

[29] Lawrence Dowler quotes Richard Szary saying that provenance information and content indexing are each half of an approach to improving description and retrieval: "The Role of Use in Defining Archival Practice and Principles," *American Archivist* 51 (winter/spring 1988): 83.

[30] Wendy M. Duff and C.A. Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 472-496.

[31] NWDA, "Executive Summary: Usability Testing Round 4" (March 12, 2008), 6, 7.

[32] E-mail correspondence 12 November 2008.

[33] Here more than one answer was possible, hence a total of over 100%. Susan Hamburger, "How Researchers Search for Manuscript and Archival Collections," *Journal of Archival Organization* 2, no. 1/2 (2004): 84.

[34] Christopher Prom, "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 67, no. 2, (fall/winter 2004): 263.

[35] Helen Tibbo and Lokman I. Meho, "Finding Finding Aids on the World Wide Web," *American Archivist* 64, no. 1 (spring/summer): 61.

[36] California Digital Library, *OAC Usability Test Summary* (October 2001), 1.

[37] Northwest Digital Archives, "Executive Summary: Usability Test #5" (November 6, 2008), 2. Published online at: http://www.orbiscascade.org/index/usability-design-working-group-reports.

[38] Kristina Southwell, "How Researchers Learn of Manuscript Resources in Western History Collections," *Archival Issues* 26, no. 2 (2002):99.

[39] Susan Hamburger, "How Researchers Search for Manuscript and Archival Collections," *Journal of Archival Organization* 2, no. 1/2 (2004): 91.

[40] Carole L. Palmer, Lauren C. Teffeau and Carrie M. Pirmann, *Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development* (Dublin, Ohio: OCLC Research, 2009): 10. Published online at: http://www.oclc.org/programs/publications/reports/2009-02.pdf.

[41] Jihyun Kim, "EAD Encoding and Display: A Content Analysis," *Journal of Archival Organization* 2, no. 3 (2004): 50.

[42] Wendy Duff and Catherine Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 493-494.

[43] Neither users nor testers ever suggest that Library of Congress Subject Headings (LCSH) are what they mean by "subjects;" LCSH is never even mentioned by testers or users.

[44] Jane Stevenson, "'What Happens If I Click on This?': Experiences of the Archives Hub," *Ariadne* 57 (October 2008). Published online at: http://www.ariadne.ac.uk/issue57/stevenson/.

[45] Wendy Duff and Catherine Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 485.

[46] Sara Snyder et al., *[Archives of American Art] Web Usability Study Report: Round 1, Subject Experts,* (April 2008).

[47] Dana Chisnell, *Report: RLG Archival Resources Service Redesign Focus Groups* (Usability Works, 2 December 2004), 14-15.

[48] Richard Lytle, "Intellectual Access to Archives: I. Provenance and Content Indexing Methods of Subject Retrieval," *American Archivist* 43 (winter 1980): 68.

[49] Elsie T. Freeman, "In the Eye of the Beholder: Archives Administration from the User's Point of View," *American Archivist* 47, no. 2 (spring 1984): 116. Wendy Duff and Catherine Johnson added the requirement that systems be intuitive: "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 472.

[50] "[R]ecords in an archives are not records *about* an activity, but records *of* an activity."[Emphasis added.] For a concise summary of how historians use provenance information for subject queries, see Wendy Duff and Catherine A. Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 477, including footnote 5.

[51] Avra Michelson, "Description and Reference in the Age of Automation," *American Archivist* 50 (spring 1987): 199.

[52] Joanne Archer, Ann M. Hanlon and Jennie Levine, "Investigating Primary Source Literacy," forthcoming [in *Journal of Academic Librarianship* 35, no. 5 (September 2009)]: 13. Archer, Hanlon and Levine report students' searches, using keywords, returned thousands of results and very few students refined or limited their results.

[53] Christopher Prom, "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 67, no. 2, (fall/winter 2004): 254.

[54] Andrea Rosenbusch, "Are Our Users Being Served? A Report on Online Archival Databases," *Archives and Manuscripts* 29, no. 1 (2001): 50.

[55] Sara Snyder et al., *[Archives of American Art] Web Usability Study Report: Round 1, Subject Experts* (April 2008).

[56] Jane Stevenson, "'What Happens If I Click on This?': Experiences of the Archives Hub," *Ariadne* 57 (October 2008). Published online at: http://www.ariadne.ac.uk/issue57/stevenson/.

[57] Dana Chisnell, *ArchiveGrid Rapid Iterative Testing Report* (Usability Works, 28 February 2006), 12, 17.

[58] Wendy Duff, and Penka Stoyanova, "Transforming the Crazy Quilt: Archival Displays from the Users Point of View," *Archivaria* 45 (spring 1998): 57.

[59] Andrea Rosenbusch, "Are Our Users Being Served? A Report on Online Archival Databases," *Archives and Manuscripts* 29, no. 1 (2001): 49.

[60] Michelle Light, "The Endangerment of Trees," EAD @ 10 (August 31, 2008): 2; *forthcoming.* Published online at: http://www.archivists.org/publications/proceedings/EAD@10/Light-EAD@10.pdf.

[61] NWDA, "Executive Summary: Usability Testing Round 4" (March 12, 2008), 4.

[62] Elizabeth Yakel et al., "Creating the Next Generation of Archival Finding Aids," *D-Lib* 13, no. 5/6 (May/June 2007). Published online at: http://www.dlib.org/dlib/may07/yakel/05yakel.html. For an intriguing example of possibilities for one specific finding aid, see Magia Ghetu Krause and Elizabeth Yakel, "Interaction in Virtual Archives: The Polar Bear Expedition Digital Collections Next Generation Finding Aid, *American Archivist* 70, no. 2 (fall/winter 2007): 282-314.

[63] Avra Michelson, "Description and Reference in the Age of Automation," *American Archivist* 50 (spring 1987): 192, 200.

[64] Karen Markey, "Twenty-Five Years of End-User Searching, Part 2: Future Research Directions," *Journal of the American Society for Information Science and Technology,* 58 no. 8 (2007): 1126.

[65] An outstanding counter-example is Jody DeRidder's use of sitemaps and links in static html pages to increase dramatically discovery of finding aids at the University of Tennessee: "Googlizing a Digital Library," *The Code4Lib Journal* 2 (March 24, 2008). Published online at: http://journal.code4lib.org/articles/43/.

[66] Cathy De Rosa et al., *College Students' Perceptions of Libraries and Information Resources* (Dublin, Ohio: OCLC, 2006), 1-1. Published online at: http://www.oclc.org/reports/pdfs/studentperceptions.pdf.

[67] Joanne Archer, Ann M. Hanlon and Jennie Levine, "Investigating Primary Source Literacy," forthcoming [in *Journal of Academic Librarianship* 35, no. 5 (September 2009)]: 11.

[68] Christopher Prom, "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 67, no. 2, (fall/winter 2004): 247.

[69] Jane Stevenson, "'What Happens If I Click on This?': Experiences of the Archives Hub," *Ariadne* 57 (October 2008). Published online at: http://www.ariadne.ac.uk/issue57/stevenson/.

[70] Erika Dowell, "Web Site Usability for Rare Book and Manuscript Libraries," *RBM* 9, no. 2 (fall 2008): 175, 179.

[71] Research Information Network, *Discovering Physical Objects: Meeting Researchers' Needs* (London, October 2008): 5. Published online at: http://www.rin.ac.uk/objects.

[72] Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *American Archivist* 68, no. 2 (fall/winter 2005): 208-263.

[73] Joshua Ranger, "Mass Digitization of Archival Manuscripts," (2008). Published online at: http://www.pacsclsurvey.org/documents/ranger/04ranger.ppt.

[74] Mark Greene, "[Draft Grant Report to NHPRC:] Basic Processing User Survey Overview," (2008).

[75] For example, RLG learned this when MARC records were mixed with archival finding aids. See Dana Chisnell, *Report: RLG Archival Resources Service Redesign Focus Groups*, (Usability Works, 2 December 2004), 18.

[76] Joanne Archer, Ann M. Hanlon and Jennie Levine, "Investigating Primary Source Literacy," forthcoming [in *Journal of Academic Librarianship* 35, no. 5 (September 2009)]: 17.

[77] Wendy Duff, and Penka Stoyanova, "Transforming the Crazy Quilt: Archival Displays from the Users Point of View," *Archivaria* 45 (spring 1998): 52-54, 59.

[78] Dana Chisnell, *ArchiveGrid Rapid Iterative Testing Report* (Usability Works, 28 February 2006), 11.

[79] Jane Lee, *OAC First round Usability Test Findings* (OAC Redesign Project, 11 September 2008), 15.

[80] NWDA, "Executive Summary: Usability Testing Round 4" (March 12, 2008), 6.

[81] Dana Chisnell, *Report: RLG Archival Resources Service Redesign Focus Groups* (Usability Works, 2 December 2004), 17.

[82] RLG in-house Archival Resources correspondence from 2003. See also Cathy De Rosa et. al., *Perceptions of Library and Information Resource* (Dublin, Ohio: OCLC Research, 2005), ix. Published online at: http://www.oclc.org/reports/pdfs/Percept_all.pdf.

[83] "Generally, participants did not talk about tools for doing research…It is unlikely that researchers approach doing research by looking for a tool for doing research." Dana Chisnell, *Report: RLG Archival Resources Service Redesign Focus Groups* (Usability Works, 2 December 2004), 13.

[84] Karen Markey, "Twenty-Five Years of End-User Searching, Part 1: Research Findings," *Journal of the American Society for Information Science and Technology,* 58 no. 8 (2007): 1071.

[85] "Le traitement des fonds demeure cependant la préoccupation des centres d'archives et le lien entre l'information décrite et la clé d'accès fournie est d'autant plus fragile dans un contexte informatique," Louise Gagnon-Arguin, "Les questions de recherché comme matériau d'études des usagers en vue du traitement des archives," *Archivaria* 46 (winter 1998): 100.

[86] Dana Chisnell, *Report: RLG Archival Resources Service Redesign Focus Groups* (Usability Works, 2 December 2004), 13.

[87] Research Information Network, *Discovering Physical Objects: Meeting Researchers' Needs* (London, October 2008), 5. Published online at: http://www.rin.ac.uk/objects.

[88] Elizabeth Yakel, "Listening to Users," *Archival Issues* 26:2 (2002): 118. Susan Hamburger, "How Researchers Search for Manuscript and Archival Collections," *Journal of Archival Organization* 2, no. 1/2 (2004): 83, 89. William J. Maher, "Use of User Studies," *Midwestern Archivist* 11 (1986): 23.

[89] Kristina Southwell, "How Researchers Learn of Manuscript Resources in Western History Collections," *Archival Issues* 26, no. 2 (2002):100.

[90] Maher, William J., "Use of User Studies," *Midwestern Archivist* 11 (1986): 22.

[91] Jane Stevenson, "'What Happens If I Click on This?': Experiences of the Archives Hub," *Ariadne* 57 (October 2008). Published online at: http://www.ariadne.ac.uk/issue57/stevenson/.

[92] Kathleen Feeney, "Retrieval of Archival Finding Aids Using World-Wide-Web Search Engines," *American Archivist* 62:2 (1999): 224.

[93] Spindler, Robert P. and Richard Pearce-Moses, "Does AMC Mean 'Archives Made Confusing? Patron Understanding of USMARC AMC Catalog Records," *American Archivist* 56 (spring 1993): 340.

[94] Michelle Light, "The Endangerment of Trees," EAD @ 10 (August 31, 2008): 3-5, *forthcoming.* Published online at: http://www.archivists.org/publications/proceedings/EAD@10/Light-EAD@10.pdf. Dryden is also concerned that we have—for years—developed standards for archival description without consulting "consumers" of these descriptions: Jean Dryden, "Do We Care What Users Want? Evaluating User Satisfaction and the LibQUAL+™ Experience," *Journal of Archival Organization* 4, no. 4 (2004): 84-85.

[95] Lorcan Dempsey says that his aphorism, "Discovery happens elsewhere," "…is influenced by Steve Rubel's use of the phrase, 'traffic happens elsewhere' in discussion of the 'Cut and Paste Web' ‹ http://www.micropersuasion.com/2007/08/the-cut-and-pas.html ›. Published online at: http://orweblog.oclc.org/archives/001430.html (September 16, 2007).

[96] William J. Maher, "Use of User Studies," *Midwestern Archivist* 11 (1986): 17.