

Social Media and Archives: A Survey of Archive Users

Bruce Washburn
Consulting Software Engineer

Ellen Eckert
Research Assistant

Merrilee Proffitt
Senior Program Officer

OCLC Research



Social Media and Archives: A Survey of Archive Users

Bruce Washburn, Ellen Eckert and Merrilee Proffitt, for OCLC Research

© 2013 OCLC Online Computer Library Center, Inc.

This work is licensed under a Creative Commons Attribution 3.0 Unported License.

<http://creativecommons.org/licenses/by/3.0/>



August 2013

OCLC Research

Dublin, Ohio 43017 USA

www.oclc.org

ISBN: 1-55653-462-0 (978-1-55653-462-1)

OCLC (WorldCat): 855368862

Please direct correspondence to:

Bruce Washburn

Consulting Software Engineer

washburb@oclc.org

Suggested citation:

Washburn, Bruce, Ellen Eckert and Merrilee Proffitt. 2013. Social Media and Archives: A Survey of Archive Users. Dublin, Ohio: OCLC Research.

<http://www.oclc.org/content/dam/research/publications/library/2013/2013-06.pdf>.

Contents

Introduction	5
Survey Methodology	5
The Role of Social Media in Archival Research	6
Key Findings	7
The utility of social media and user-contributed content depends on the credibility and expertise of the source	7
Unaffiliated scholars value archives	7
Appendix 1: Survey Questions, Summary of Responses, and Notes	9
Appendix 2: Understanding Special Collections Survey Instrument	24
Appendix 3: About ArchiveGrid.....	29
Appendix 4: Survey in Context: Prior User and Usability Studies.....	30
References	32

Figures

Figure 1: Have used special collections materials	9
Figure 2: Important attributes of special collections used	11
Figure 3: Research subjects in the last year	12
Figure 4: Intended purpose of research	13
Figure 5: Usual role when using special collections	14
Figure 6: Type of help utilized in initial investigations of research.....	15
Figure 7: Resources most useful when in the middle of, or completing research	16
Figure 8: All potential sources checked at completion of research.....	17
Figure 9: Ways of discovering new websites and other research resources.....	18
Figure 10: How sharing information about a new website or resource is identified.....	19
Figure 11: Preferred comm. when sharing information about a new website or resource	20
Figure 12: Website features considered valuable for research	21
Figure 13: Type of user whose comments, recommendations, etc. is found most valuable ...	22
Figure 14: Reasons for not using special collections	23

Introduction

In April and May of 2012, OCLC Research conducted a survey of users of archives to learn more about their habits and preferences. In particular, we focused on the roles that social media, recommendations, reviews, and other forms of user-contributed annotation play in archival research. We surveyed faculty, graduate students, and genealogists across North America and a few in the United Kingdom. Data collected from 695 respondents who completed the survey broadened our knowledge of who uses archives and special collections for research. We also learned how users communicate and share information about systems like OCLC Research's ArchiveGrid, and about the relative importance of social media tools in an archival discovery environment.

Survey Methodology

Our survey asked 15 questions, structured around four primary areas of interest:

1. Who are archival and special collections researchers, what are their roles, and what is their research objective?
2. How are research resources found, and how are those discoveries shared?
3. Where do researchers turn for support at the beginning and during the research process?
4. What role do social media play in the research process, and how important are other review, recommendation and annotation features?

Most questions were multiple choice and allowed respondents to choose more than one answer and leave comments. This mix of qualitative and quantitative data supported standard analytics while offering the opportunity for a richer and more nuanced view, in particular capturing free-text responses that our survey questions might have otherwise missed.

The survey was available on SurveyMonkey.com between 27 April and 25 May 2012.

We promoted the survey through Twitter, Facebook, listservs, mailing lists, and individual contacts. To encourage survey responses, we offered one respondent chosen at random a \$50 Amazon gift card.

The complete survey instrument can be found in appendix 2.

The Role of Social Media in Archival Research

Interest in including social media features (also known as Web 2.0) in libraries and archives (dubbed “Archives 2.0” by practitioners) has been a topic of interest, investigation, and experimentation for some time, generating numerous papers and many conference presentations. In formulating our own work, we took note of some of the more significant bodies of work in this area. Between 2009 and 2010, a 21-member OCLC Research Library Partnership working group reviewed 76 sites relevant to libraries, archives, and museums that supported social media features (summarized in three reports plus an executive summary) (OCLC Research 2012). One aspect of the study was a survey of site managers: since the survey did not include a question about analysis done prior to implementation, we don’t know how many, if any, of the sites included undertook work similar to what we’ve outlined in this paper.

Similarly, the book *A Different Kind of Web* (Theimer 2011) features 13 in-depth archival case studies focusing on social media and outreach. The case studies do not feature assessment of target audiences to evaluate the potential impact and benefit of social media implementation. And, in many cases, such an analysis may not have been appropriate since the point of many of the projects was to reach new audiences.

The Northwest Digital Archives “Researcher Needs Study” (2008-2009) came closest to looking at the issues we were most interested in. In their study, subjects were decidedly mixed or negative about the inclusion of social media features. However, subjects in this study were mostly administrators, with few faculty members or genealogists represented (the balance of subjects were students, with no differentiation between graduates and undergraduates, and “serious amateurs”).

In the case of ArchiveGrid, we had an installed user base, so to speak, and we wanted to weigh, in an informed way, the potential benefit of new features against the very real costs of implementation and maintenance. The Social Metadata group appropriately acknowledged that “We are approaching the end of the “wild west” of “Web 2.0” (Smith-Yoshimura 2011, 8-9) and that it is time for “data-driven decisions about launching, expanding, or ending our social media experiments.” Additionally, they recommend that implementation not only include user testing after implementation, but also before, in order to establish requirements.

Key Findings

The complete survey and responses are available in appendices 1 and 2. The survey confirms much of what we had understood from related work and prior studies. We note these key findings.

The utility of social media and user-contributed content depends on the credibility and expertise of the source

The survey responses indicate that e-mail and word of mouth continue to be the primary ways archival researchers share information about the resources they discover. Features such as tags, reviews, recommendations and user comments are viewed as useful by fewer than half of those responding.

The perceived utility of user-contributed content greatly increases if the researcher can effectively gauge the credibility and expertise of the source. Researchers favor the recommendations and commentary of archivists and librarians over other groups, apart from any additional understanding of the content contributor's role and expertise.

Considering the important role that e-mail plays in communicating about resources, it's important for providers of information resources to continue to invest energy and effort into communications that are tailored for that medium. Even with increasing attention shifting to communication on the web, providers should not neglect opportunities to reach users directly on e-mail lists that they may subscribe to.

Based on what we learned, we decided against including social media tools in ArchiveGrid. Instead, we have begun exploring how archivists and librarians might provide information on topical relations across collections.

Unaffiliated scholars value archives

When we designed ArchiveGrid we had three primary audiences in mind: faculty, graduate students, and genealogists. The survey responses reflect these roles, and also show some overlap, with respondents selecting more than one role. For example: Twenty-seven of the 208 respondents (13%) who identified themselves as faculty also identified themselves as graduate students.

But nearly one-quarter of all survey respondents identified themselves as "unaffiliated scholars" when describing their role while conducting research in archives and special collections. We define an unaffiliated scholar as one not doing research for academic or genealogy purposes.

The roles this group provided in the survey cover a wide range of disciplines and interests:

- Writers and authors
- Teachers
- Filmmakers
- Bloggers
- Lecturers
- Exhibit makers

And the survey responses suggest these general characteristics for this group:

- Use “Web 2.0” tools
- Computer literate
- Collaborate online
- Like to learn about new resources
- Not bound to the institutions many archives and special collections are affiliated with
- Still need librarians and archivists for research help

The purpose for their research is generally divided between personal interest (70%) and publication (72%). Publication interests are not limited to academic journals. They share information about their research in ways similar to faculty, graduate students, and genealogists, with e-mail being their primary method of sharing.

Appendix 1: Survey Questions, Summary of Responses, and Notes

1. Have you used special collections materials?

(Special collections materials are defined as library and archival materials in any format, generally characterized by their value, physical format, uniqueness or rarity. For example: rare books, manuscripts, photographs, institutional archives including digital items).

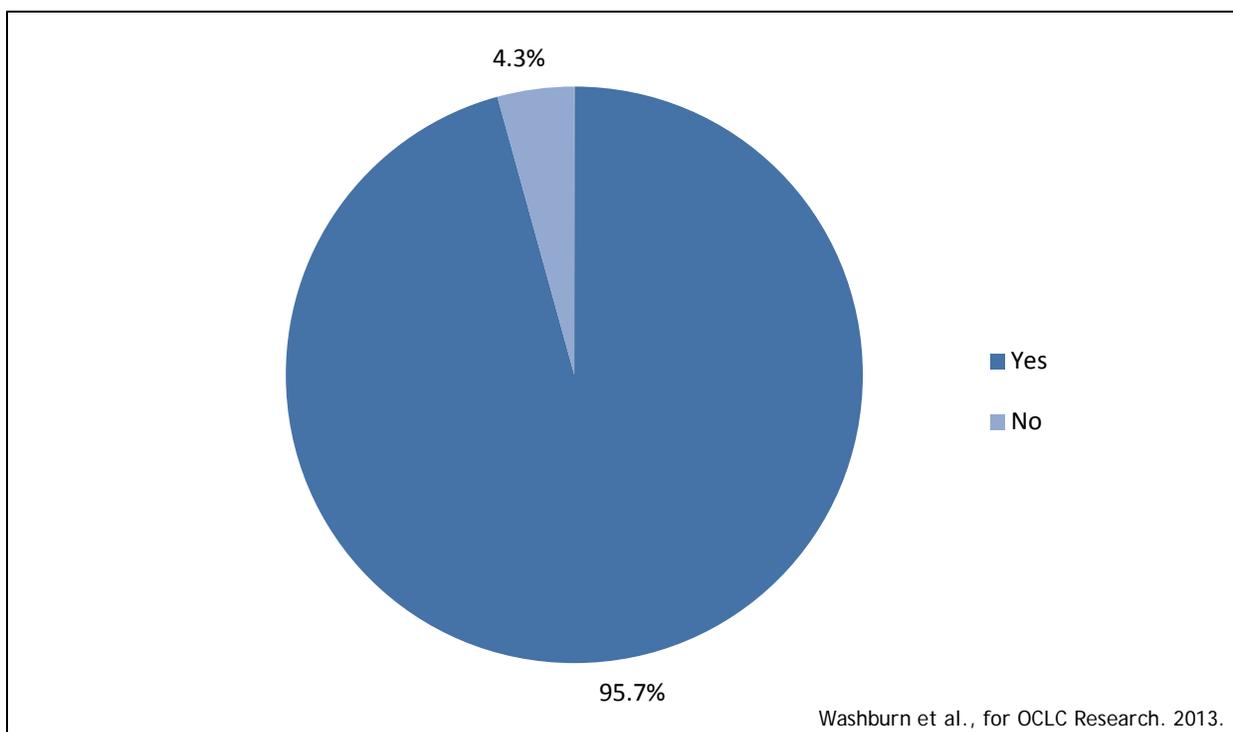


Figure 1: Have used special collections materials

Your Use of Special Collections

2. What kind of special collections materials did you use?

This was a free text response and the following excerpt is typical of the primary source materials described:

- Early genealogies of Virginia families
- Correspondence
- Letters, diaries, manuscripts, printed material from the 18th century, old maps, photos
- Maps
- Letters, notes, newspaper clippings
- Manuscripts and early printed books
- Manuscripts
- Letters, probate files, business documents
- French government records; Catholic Church records; U.S. Federal government records
- Personal papers, photo collections
- Rare books; manuscript materials; engravings/prints
- Manuscripts, records, personal papers, magazines and newspapers, visual material
- I have used a variety of special collections materials. Mostly archival material (manuscripts, papers, photographs, institutional records...) but have also used rare books
- Personal papers, consular archives, court records (all 19th century)
- Items in "rare book" rooms, non-circulating specialized library collections, archival material

A relatively small number of responses were provided in terms of specific collections or online discovery systems, rather than material types:

- The US Coast Guard Auxiliary Special Collection
- Memoir of Brass Crosby
- Record Group 56 at the National Archives
- Duke University archives digital collections online
- Online resources at the Archives of American Art

3. What are the important attributes of these materials for you?

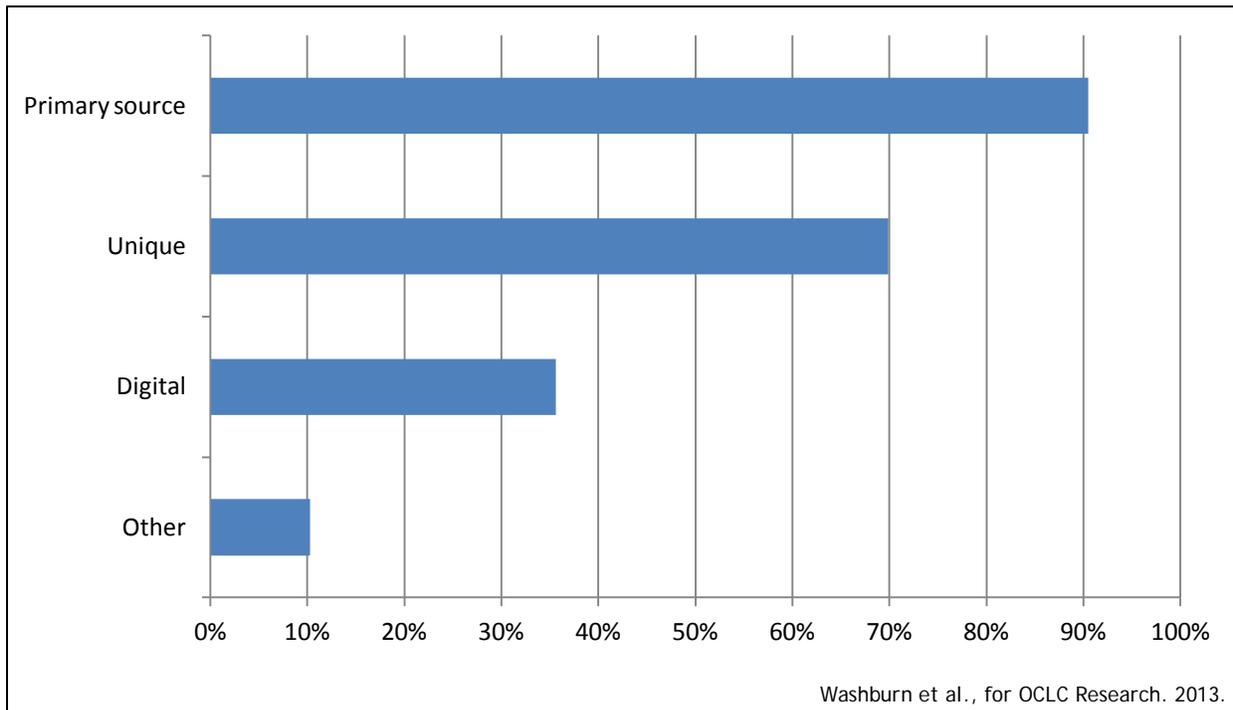


Figure 2: Important attributes of special collections used

Free text response highlights:

- Unpublished and never before cited
- If they had been digitized, it would have been helpful
- Rare book had a pull out map of the town my ancestors established with their names listed on it. Holding it in my hands & perusing the “real thing” was enchanting...so much more so, than seeing it online.

Research Areas

4. In the last year or so, what have been the subjects of your research?

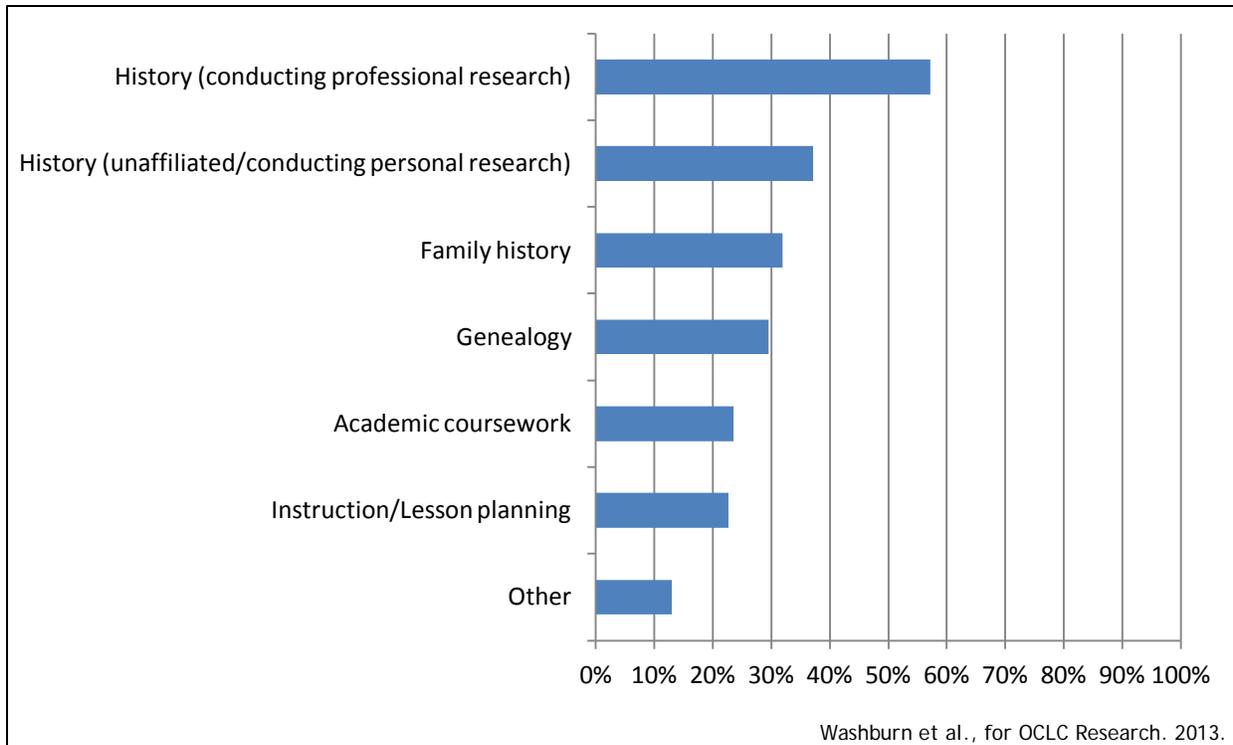


Figure 3: Research subjects in the last year

Free text response highlights:

- Historical research in support of archival appraisal and large artifact preservation reporting.
- Inspiration for creative work
- Provenance research (conducting professional research)
- Curiosity
- Art, technology, science, sewage, bikes, barefoot running, chicken reproduction, home brewing, etc.

Research Purposes

5. What is the intended purpose of your research?

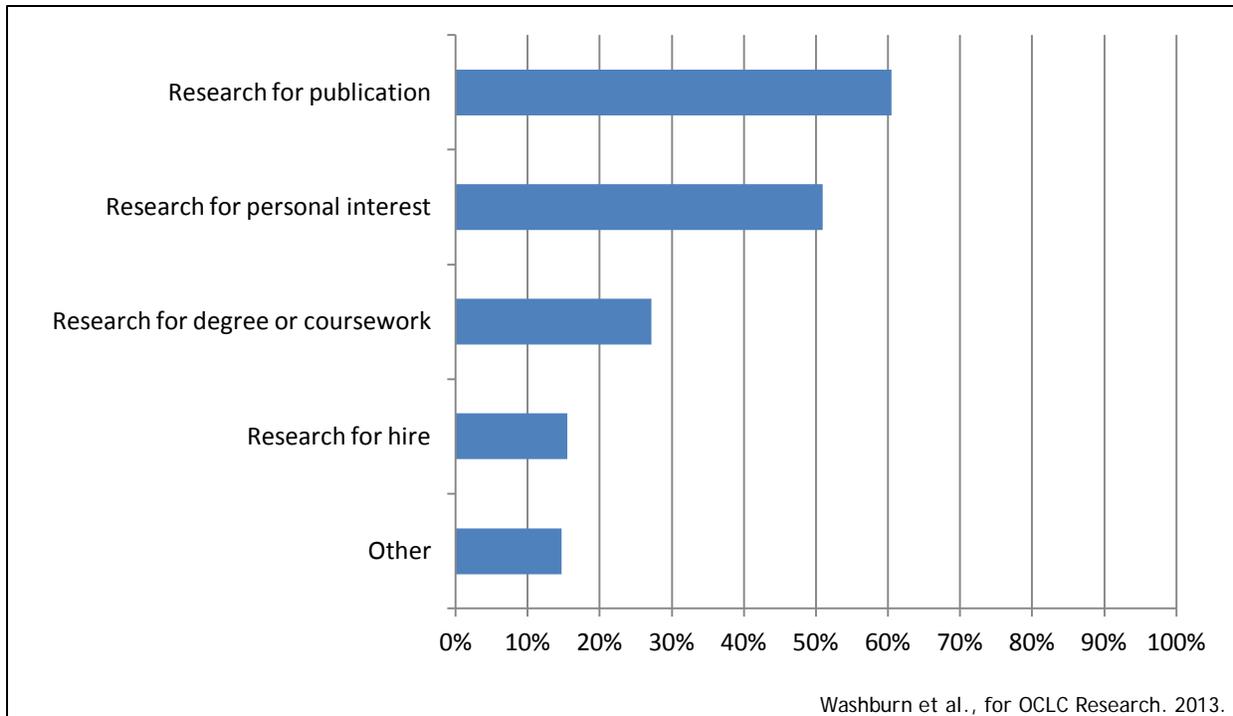


Figure 4: Intended purpose of research

Free text response highlights:

- Documentation, authentication, “fleshing out” individuals, broadening my own understanding
- “Publication” of other kinds, such as conference presentation and website publication
- Push back the frontiers of ignorance
- Reference archivist—I conduct research on behalf of off-site patrons
- Compiling a Family History to distribute to Family members.

Your Role as a Researcher

6. When using special collections what is your usual role

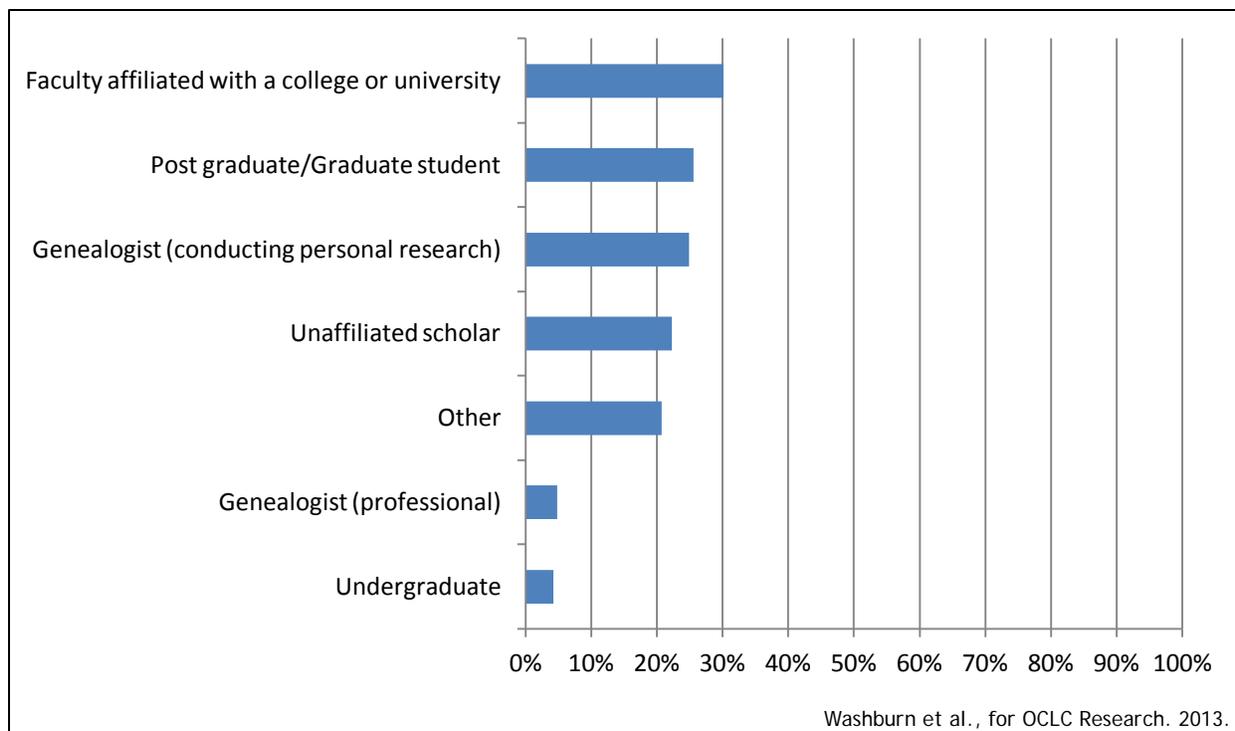


Figure 5: Usual role when using special collections

Free text response highlights:

- Pensioner botanist
- In my role as project manager for a research institute
- Curator and PhD student
- Local history buff
- Most recently, I was working for an academic library (but using the special collections at another library to do historical/background research for a project).

The Research Process

7. Remembering your research in the past year or so, as you begin the research process where do you typically go for help in your initial investigations?

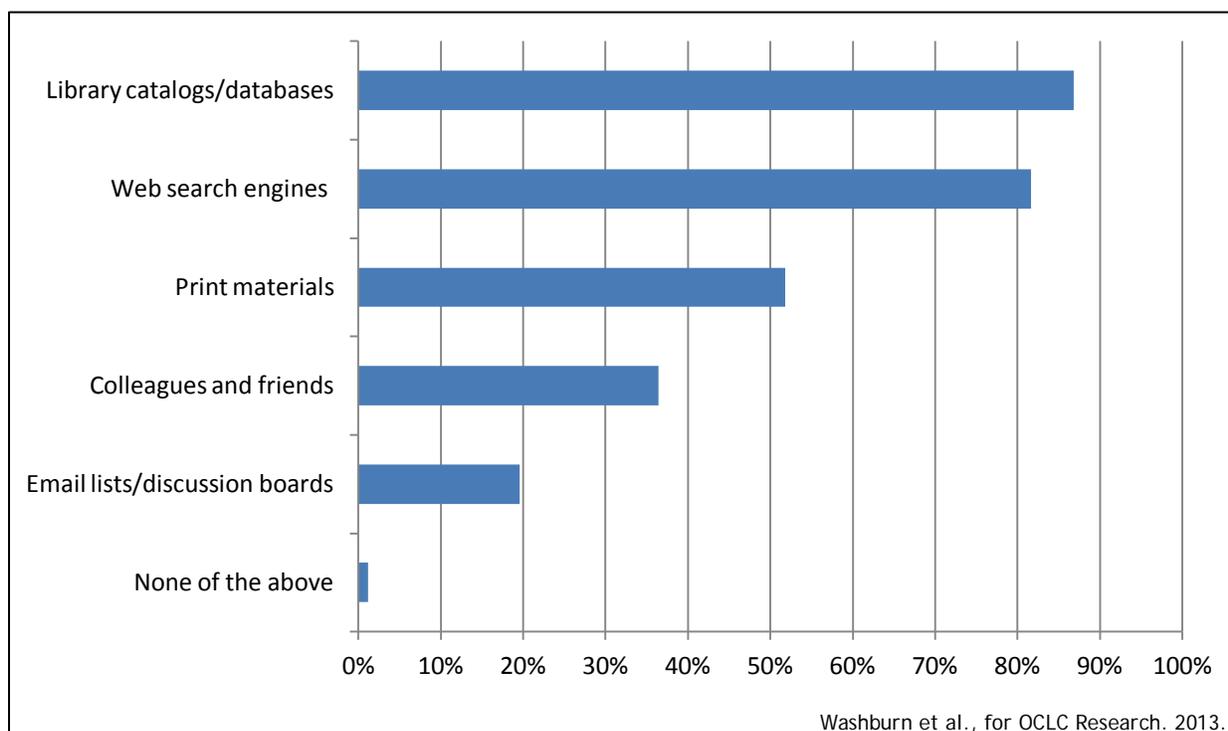


Figure 6: Type of help utilized in initial investigations of research

Free text response highlights:

- Relatives—especially my mother and other older relatives, who have memories about certain ancestors and family events.
- WorldCat, finding aids, archive.org, Google News archive, Google Books, Google Scholar, HathiTrust, etc.
- Ancestry.com; Numerous aviation historical websites
- First, my personal 3x5 card file grown over about 70 yrs, next, personal library reference books (in which many loose notes/articles are filed, next the internet.
- eBay, Amazon, the Internet Archive, and rare book dealer sites (not to purchase but to find out what sources exist)

The responses to question 7 conflict with other recent analyses which suggest that research typically begins in web search engines. Our allowing the selection of more than one category and our combination of library catalogs and databases in the same selection may have contributed to the unexpectedly high response rate for library catalogs/databases.

8. When you are in the middle of, or completing, your research, which resources are the most useful to you?

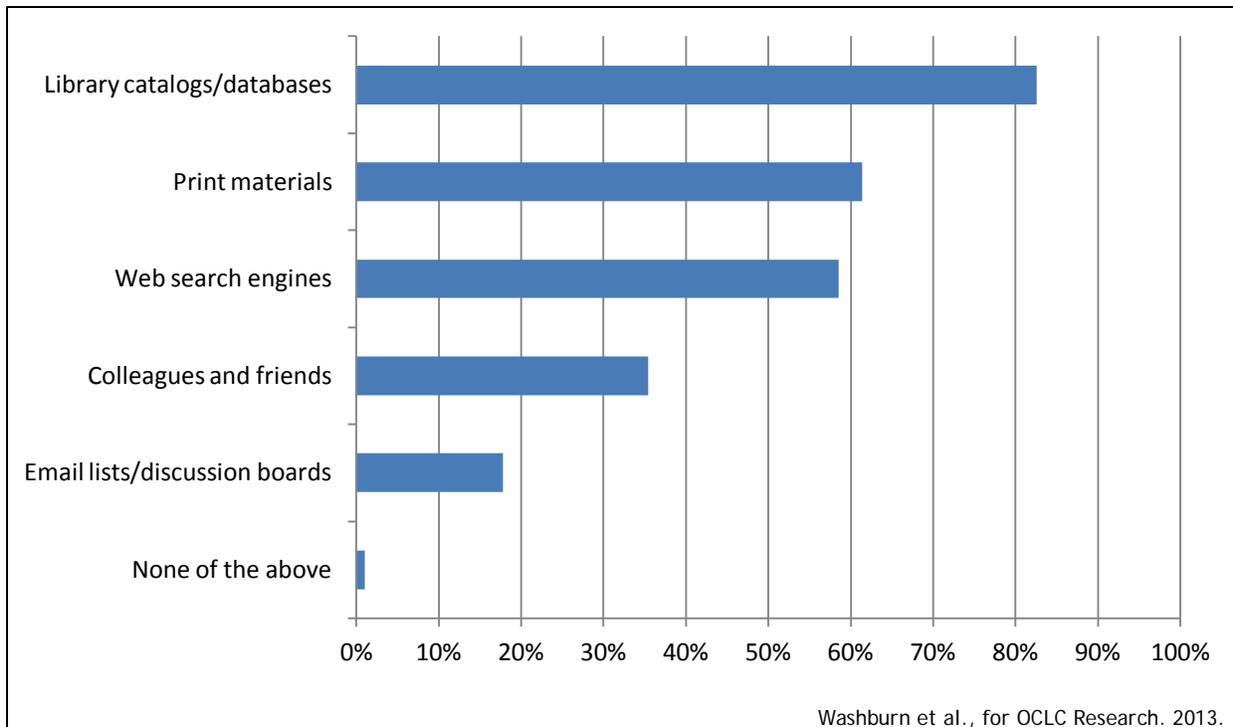


Figure 7: Resources most useful when in the middle of, or completing research

Free text response highlights:

- As I complete: web search engines less for research than for confirming biblio data.
- Site visits to appropriate archives or locations
- Digital archives like (depending on the topic) OAC, Densho, UNC’s SHC, etc.
- Archivists - Not all small collections are well-catalogued; Finding Aids
- I write to librarians and archivists. I use NUCMC or whatever it’s called now.

The responses to question 8 conflict with other recent analyses which suggest that research typically begins in web search engines. Our allowing the selection of more than one category and our combination of library catalogs and databases in the same selection may have contributed to the unexpectedly high response rate for library catalogs/databases.

9. When you complete your research, do you need to make sure that all potential sources have been checked?

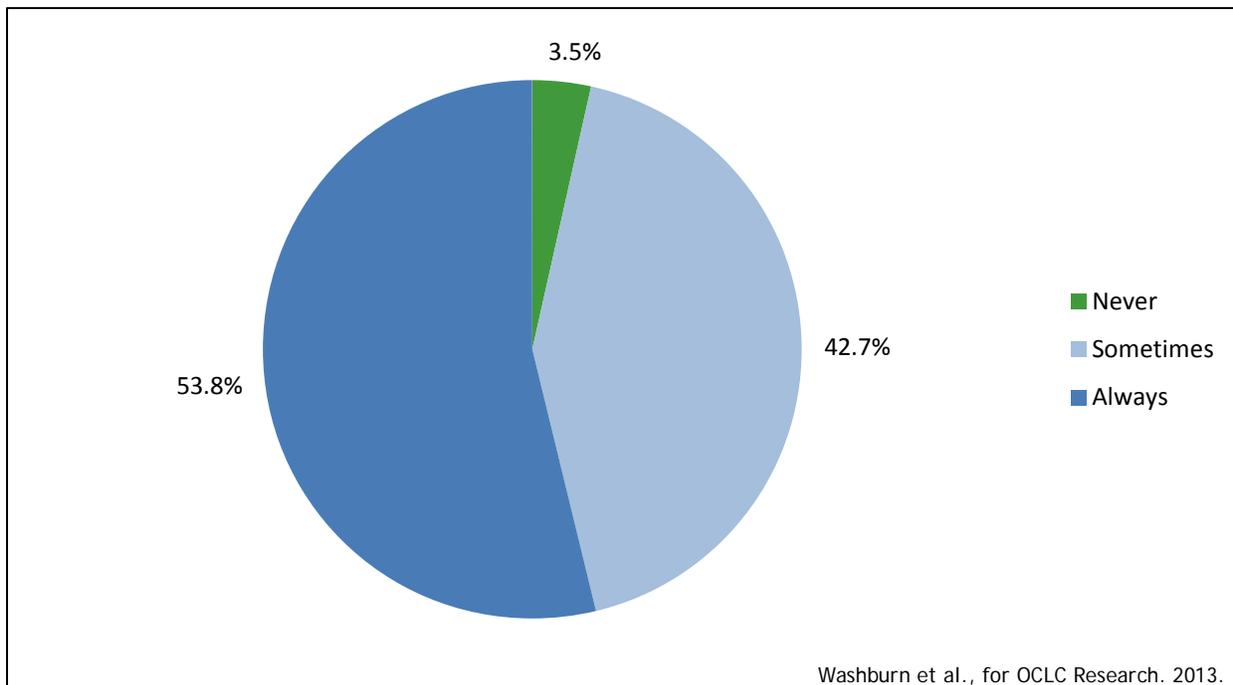


Figure 8: All potential sources checked at completion of research

Discovering New Websites and Other Research Resources

10. How do you discover new websites and other research resources?

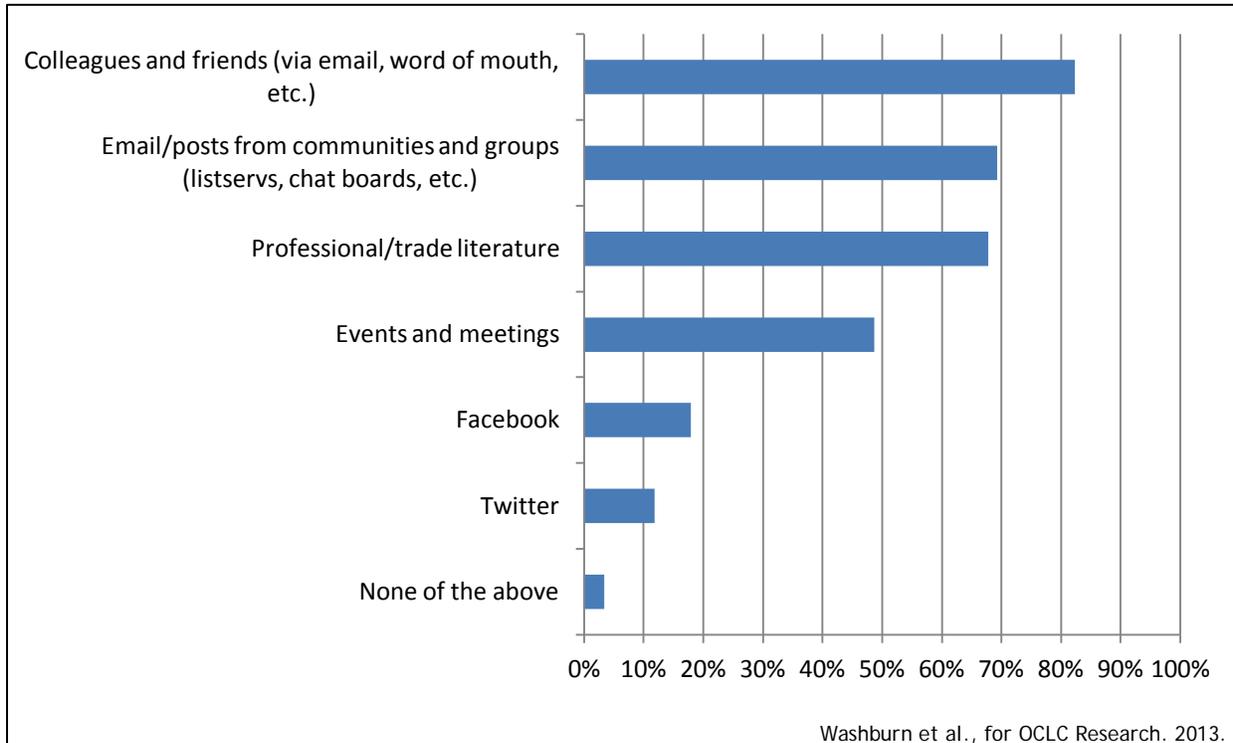


Figure 9: Ways of discovering new websites and other research resources

Free text response highlights:

- Chats with archivists and librarians
- Find them during a Google search
- The librarian at my university affiliated with the history department sends out emails about new databases and resources.
- I check every six months or so for new sources using an ordinary search engine
- I spend a lot of time on my own, prowling the web. This is my best strategy.

11. When you want to share information about a new website or research resource, how do you usually identify it?

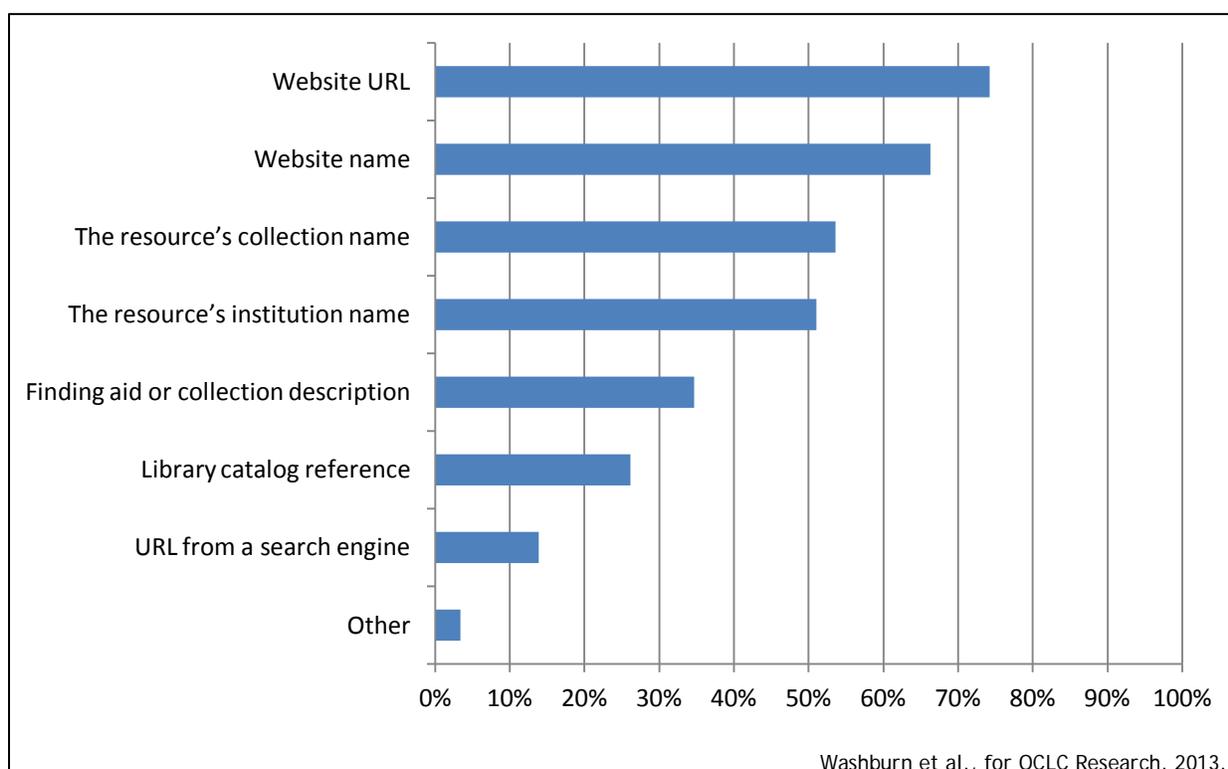


Figure 10: How sharing information about a new website or resource is identified

Free text response highlights:

- Bring it to someone's attention so they can Google it
- Whatever Chicago or APA style requires.
- I sometimes check what happens if you search for the title on Google: ideally the website will be the first hit
- I have had no occasion to share yet.
- I embed URL in website name in email

12. When you want to share information about a new website or other research resources what are your preferred ways to communicate?

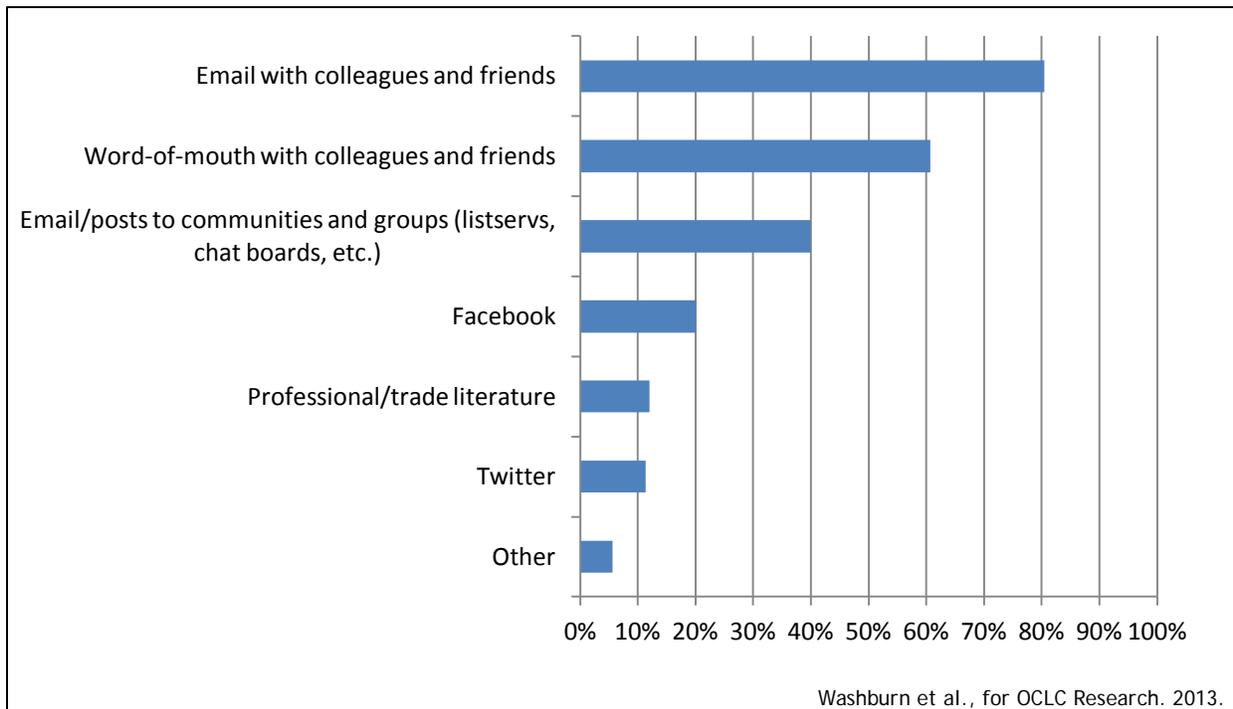


Figure 11: Preferred communication when sharing information about a new website or resource

Free text response highlights:

- Course syllabi, blackboard
- The telephone is still great
- I co-author a blog that is partially devoted to sharing resources
- Diigo set to provide weekly post to my blog
- Article in church newsletter

User Comments, Reviews, Recommendations, Tags, and Other Website Features

13. Which of these website features are valuable for your research?

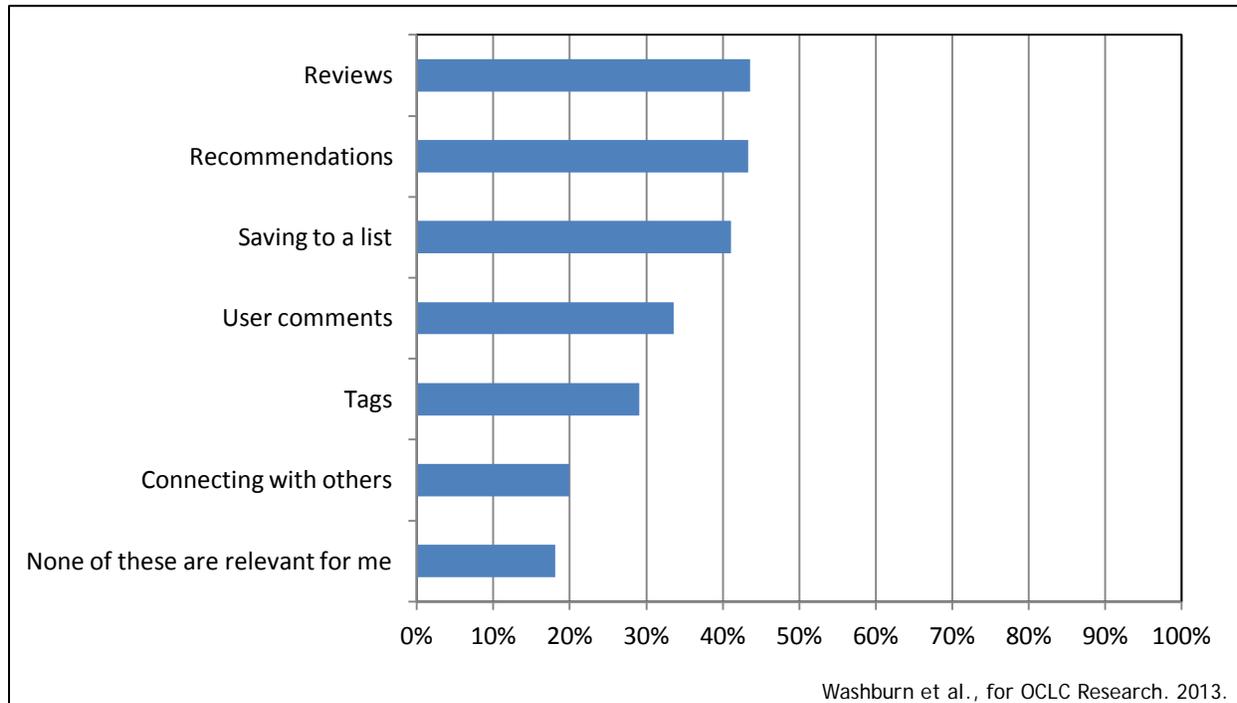


Figure 12: Website features considered valuable for research

Free text response highlights:

- I've not seen much in the way of user comments, recommendations, or reviews... sounds interesting.
- I also really like a preview DESCRIPTION so I don't waste my time if it isn't what I want.
- I don't think any of the above are "very valuable"—I find all the above useful
- I once managed to unravel a whole series of problems thanks to an anonymous comment on a library's flickrstream. I think crowdsourcing archival research is an amazing thing
- Saving to a list makes my research more efficient; user comments may be useful but often are not

14. Comments, tags, reviews and recommendations can come from a variety of sources. Is there a particular type of user whose comments, recommendations, etc. you find most valuable?

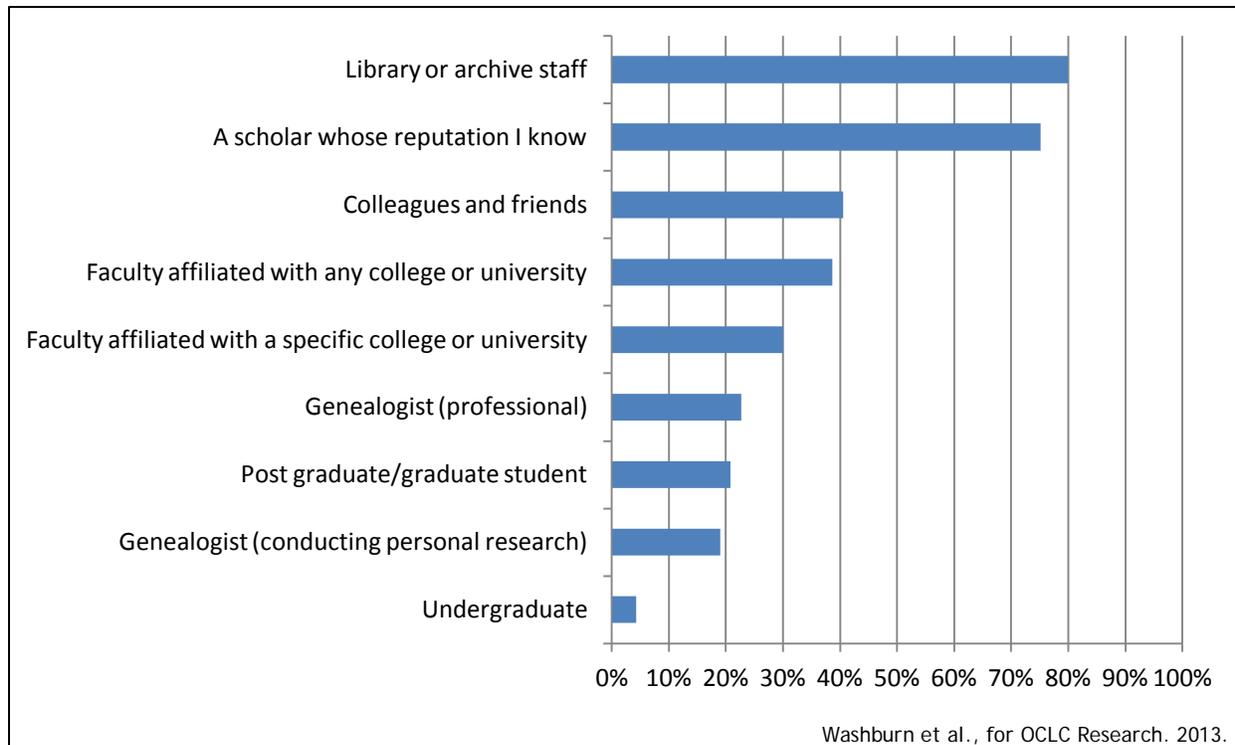


Figure 13: Type of user whose comments, recommendations, etc. is found most valuable

Free text response highlights:

- Distant relatives: their facts may not be exactly right, but they open new avenues of thought.
- I've never seen anyone leave comments or tags
- Well-written comments, etc. carry weight with me.
- Very, very few comments/tags/reviews/recommendations are worthwhile. They must be from impeccable, incredibly high-caliber sources to be worth attention.
- Doesn't matter—as long as the content is good and tags are accurate

Special Collections Use

15. If you didn't use special collections materials, was it because you:

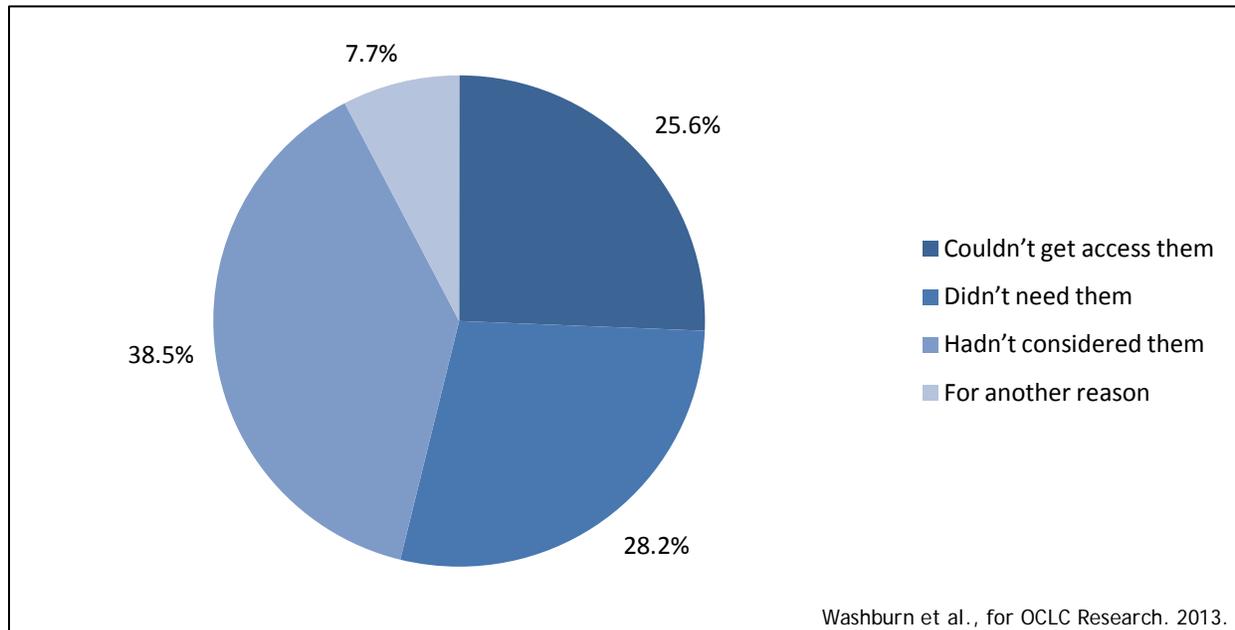


Figure 14: Reasons for not using special collections

Appendix 2: Understanding Special Collections Survey Instrument

1. Have you used special collections materials?

Special collections materials are defined as library and archival materials in any format, generally characterized by their value, physical format, uniqueness or rarity. For example: rare books, manuscripts, photographs, institutional archives including digital items.

[If the response is “no”, go to question 15.]

2. What kind of special collections materials did you use?

3. What are the important attributes of these materials for you?

- Unique
- Primary Source
- Digital
- Other

4. In the last year or so, what have been the subjects of your research?

- Family History
- Genealogy
- History (unaffiliated/conducting personal research)
- History (conducting professional research)
- Academic Coursework
- Instruction/Lesson Planning
- Other

5. What is the intended purpose of your research?
 - For publication
 - For degree or coursework
 - For hire
 - For personal interest
 - Other

6. When using special collections what is your usual role?
 - Faculty affiliated with a college or university
 - Post graduate/Graduate student
 - Undergraduate
 - Unaffiliated Scholar
 - Genealogist (professional)
 - Genealogist (conducting personal research)

7. Remembering your research in the past year or so, as you begin the research process where do you typically go for help in your initial investigations?
 - Web search engines
 - Library catalogs/databases
 - Colleagues and friends
 - E-mail lists/discussion boards
 - Print materials
 - None of the above

8. When you are in the middle of, or completing, your research, which resources are the most useful to you?
 - Web search engines
 - Library catalogs/databases
 - Colleagues and friends

- E-mail lists/discussion boards
- Print materials
- None of the above

9. When you complete your research, do you need to make sure that all potential sources have been checked?

- Never
- Sometimes
- Always

10. How do you discover new websites and other research resources?

- Colleagues and friends (via e-mail, word of mouth, etc.)
- Professional/trade literature
- Events and meetings
- E-mail/posts from communities and groups (listservs, chat boards, etc.)
- Twitter
- Facebook
- None of the above

11. When you want to share information about a new website or research resource, how do you usually identify it?

- Website name
- Website URL
- URL from a search engine
- The resource's institution name
- The resource's collection name
- Finding aid or collection description
- Library catalog reference
- Other

12. When you want to share information about a new website or other research resources what are your preferred ways to communicate?

- Word of mouth with colleagues and friends
- Professional/trade literature
- E-mail with colleagues and friends
- E-mail/posts to communities and groups (listservs, chat boards, etc.)
- Twitter
- Facebook
- Other

13. Which of these website features are valuable for your research?

- User comments
- Tags
- Reviews
- Recommendations
- Saving to a list
- Connecting with others
- None of these are relevant for me

14. Comments, tags, reviews and recommendations can come from a variety of sources. Is there a particular type of user whose comments, recommendations, etc. you find most valuable?

- A scholar whose reputation I know
- Faculty affiliated with any college or university
- Faculty affiliated with a specific college or university
- Library or archive staff
- Undergraduate
- Post graduate/graduate student
- Genealogist (professional)

- Genealogist (conducting personal research)
- Colleagues and friends

15. If you didn't use special collections, was it because you:

- Didn't need them
- Couldn't access them
- Hadn't considered them

Appendix 3: About ArchiveGrid

ArchiveGrid is a collection of nearly two million archival material descriptions, including MARC records from WorldCat and finding aids harvested from the web. It's supported by OCLC Research as the basis for our experimentation and testing in text mining, data analysis, and discovery system applications and interfaces. Archival collections held by thousands of libraries, museums, historical societies, and archives are represented in ArchiveGrid.

ArchiveGrid provides access to detailed archival collection descriptions, making information available about historical documents, personal papers, family histories, and other archival materials. It also provides contact information for the institutions where the collections are kept.

ArchiveGrid data is primarily focused on archival material descriptions for institutions in the United States. This reflects the contribution patterns for descriptions of materials under archival control in WorldCat, which make up the majority of descriptions in ArchiveGrid. We may extend ArchiveGrid beyond its current scope if it is necessary to support OCLC Research experimental objectives.

ArchiveGrid illustrates OCLC's interest in advancing issues important to the archival community. Our work within ArchiveGrid gives OCLC Research a foundation for collaboration and interactions with others in the archival community. We expect to share the results of MARC and EAD tag analysis, provide discovery system analytics for contributors, document investigations of text mining and data visualization, participate in community working groups pursuing improvements to description and discovery, and more. To support those interests and objectives, we'll continue to build this extensive and current aggregation of archival material descriptions, within the constraints of OCLC Research's committed and on-going support for this project.

OCLC had offered ArchiveGrid as a subscription-based discovery service until 2012 when that subscription service was discontinued. While the new, freely-available OCLC Research ArchiveGrid interface is not a full production service, it shares some of the same attributes. Researchers can expect to use it for discovery of archival materials, and archives can work with OCLC Research to have their materials represented in the aggregation in a reliable and persistent way.

Appendix 4: Survey in Context: Prior User and Usability Studies

ArchiveGrid developed from an RLG archival discovery system, “RLG Archival Resources.” In 2004, with the help of a consultant, Dana Chisnell, RLG conducted three focus groups to gain a better understanding of the potential users of RLG Archival Resources, which RLG had defined as faculty, graduate students, and genealogists. Nineteen participants representing each user group were asked to interact with a prototype. These focus groups were “similar to typical marketing focus groups in that we were able to gather qualitative data about user preferences, but here the format also provided opportunities to hear about the types of work people do and how the service might fit into that.” (Chisnell 2004).

Four research questions scoped the 2004 study:

1. Who are ArchiveGrid users and what are their characteristics?
2. What resources do researchers value and how do they find them?
3. How do researchers search online?
4. What do researchers want to know from their search?

In 2006, RLG conducted usability studies with 20 faculty, graduate students, and genealogists—the three previously identified user communities—using a functional prototype of ArchiveGrid. (Chisnell 2006)

Since these earlier user studies had been conducted prior to the widespread availability of annotation and review and recommendation features on the web, the 2012 survey was chiefly designed to learn how ArchiveGrid users would approach social media tools, but we also saw it as an opportunity to refresh our notions of user populations from the 2004 and 2006 studies.

Previous user studies had relied on an outside contractor, and had been costly and labor-intensive in terms of identifying and recruiting appropriate focus group and user study participants, as well as intensive preparation and participation by staff. With our 2012 survey, we opted to conduct a web survey using SurveyMonkey. Recognizing that an online survey would not provide the ability to interact with participants in the same way (pose

follow up questions, ask for clarifications, or gather reactions to wireframes, for example), the team opted for the online survey as a “good enough” solution that would allow us to get valuable information that would inform development planning and allow for wise use of limited resources.

In preparation for our survey, we conducted a review of other relevant studies, paying particularly close attention to the findings of the Northwest Digital Archives Digital Program Needs Assessment report published by Jodi Allison-Bunnell, Elizabeth Yakel, and Janet Hauck (2009) and the reports and presentations associated with the Online Archive of California’s OAC Redesign Project (CDL 2013).

References

Allison-Bunnell, Jodi, Elizabeth Yakel and Janet Hauck. 2009. *Digital Program Planning: Researcher Needs Study Report on Findings*. Eugene, Oregon: Orbis Cascade Alliance Northwest Digital Archives http://www.orbiscascade.org/index/cms-file-system-action/nwda/files/researcher_needs_study_report_final_20090722.pdf.

CDL (California Digital Library). 2013. "OAC Redesign Project." http://www.cdlib.org/services/dsc/projects/oac_redesign.html.

Chisnell, Dana. "RLG Archival Resources Service Redesign Focus Groups." Commissioned by Research Libraries Group. Unpublished report, last modified December 3, 2004. Microsoft Word file.

Chisnell, Dana. "ArchiveGrid Rapid Iterative Testing." Commissioned by Research Libraries Group. Unpublished report, last modified February 28, 2006. Microsoft Word file.

OCLC Research. 2012. "Social Metadata for Libraries, Archives, and Museums." <http://www.oclc.org/research/publications/library/2012/2012-01r.html>.

Smith-Yoshimura, Karen. 2011. *Social Metadata for Libraries, Archives, and Museums: Executive summary*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2012/2012-02.pdf>.

Theimer, Kate. 2011. *A Different Kind of Web: New Connections Between Archives and Our Users*. Chicago, Illinois: Society of American Archivists.