Kris Bronstad

# References to Archival Materials in Scholarly History Monographs

*This study looks at citations of archival material in a sample of 136 recently published scholarly historical monographs produced by a selection of highly cited university presses in the United States, with the goal of discovering patterns in scholarly user reportage of archival use. The study found that 68 percent of the titles referenced at least one archival collection, that archival collections housed at universities were used more often than other types of repositories, and that the amount and type of repositories did not in most cases vary based on the subject matter of the book. The study also revealed that less than 3 percent of all archival citations in the books examined were to digital collections. The findings could potentially provide a baseline by which further and more diverse archival use and users can be measured.*

Citation analysis, or the counting of the number of times specific resources are cited in scholarly publications, is a common tool in library and information sciences and has been in practice for around a century. This form of analysis is popular because it provides metric-based feedback about use of documents. But it has rarely been used to investigate use of archival material.

I have attempted to adapt citation analysis to the examination of archival collection and repository use. In my study, I have set out to look at citations of archival material in a sample of 136 recently published, single-authored books (or monographs) produced by highly cited university presses in the United States and made available as e-books and categorized according to subject by Project Muse. The intended goal is to examine the feasibility of using citation analysis to discover patterns in scholarly user reportage of archival use. Factors studied include the percentage of monographs that cite archival material; variations in number of formats (physical or digital) and repositories referenced; and whether specific historical subjects are more likely to follow distinct usage patterns.

It should be noted that unlike a strictly defined "citation study," which counts each time a source is cited, this is a "reference study," which counts each archival collection only once per book, even if it is cited multiple times.

## Research Problem

Understanding archival use has long been considered a vital issue to archivists, and user studies employing a number of methodologies have often been called for and periodically undertaken. Although citation or reference studies are not often considered in archival literature, fields that produce published works put considerable weight behind tracking and counting citations. The sources of published work in these fields (including scholarly articles in the field of history) are easily measured and tracked using citation tools such as Google Scholar, Scopus, or Web of Science. Archival citations, however, are not available through such indexes. That absence prohibits archivists and scholars from seeing a broader network of archival collections' roles in scholarly communication. Without that broader understanding, archivists lack access to existing evidence that may prove helpful in evaluating larger trends, movements, and impacts that affect the field as a whole.

To that end, this project intended to answer the question: What can references to archival material in scholarly historical monographs tell us about recent scholarly use of archival material? More specifically, can we tell from archival references what historical specialties are most likely to use archival material, and can we observe any patterns in how they may use different formats or repositories?

## Literature Review

Evaluating the number and nature of citations in published materials to other works as a way to understand use has been an established practice in library and information studies, particularly when reviewing the "hard sciences." There is a thinner history of citation studies concerning the humanities[1] and many citation studies of humanities scholars themselves exclude the study of archival sources.[2]

There are exceptions: Some citation studies investigating the information use of historians by study of references in scholarly history journals have included information about archival sources. Jones, Chapman, and Woods looked at 7,000 references in historical publications and found that 12.6 percent of all references were to unpublished material, and 55 percent of all references to manuscripts came from the British Museum of Public Reference Office.[3] Richard Heinzkill analyzed 9,556 footnotes in 14 journals and found that less than 5 percent of the references were to manuscripts.[4] Clark Elliot, investigating material in the history of science, found

---

1. Richard Heinzkill, "Characteristics of References in Selected Scholarly English Literary Journals," *Library Quarterly* 50, no. 3 (1980): 352–65.
2. For one recent example, see Jennifer E. Knievel and Charlene Kellsey, "Citation Analysis for Collection Development: A Comparative Study of Eight Humanities Fields," *Library Quarterly* 75, no. 2 (2005): 142–68.
3. Clyve Jones, Michael Chapman, and Pamela Carr Woods, "The Characteristics of the Literature Used by Historians," *Journal of Librarianship and Information Science* 4, no. 3 (1972): 137–56.
4. Heinzkill, "Characteristics of References in Selected Scholarly English Literary Journals."

that 28 percent of 3,600 references in 15 journals were to unpublished sources.[5] But most relevant to this study is the work of Frederic Miller, who looked at archival use in 214 journals, tabulating how many repositories, collections, and series were used.[6] While similar, my study differs from Miller's in two significant ways: Miller's is a true citation study, interested in looking at "significance of use" of a collection or series as reflected in how many citations there are to a particular collection.[7] Second, I do not look at individual series or collections, as Miller does in his work; instead, I merely count each time any collection is used from the same repository in the same book.

Other reference studies are limited to following past researchers of certain repositories. In 1986, Jacqueline Goggin examined the works of users of thirteen black and women's organization collections housed at the Library of Congress.[8] Like Goggin, Nancy McCall and Lisa A. Mix used researcher records to track citations, and therefore use, of papers; their study looked at the cited use of the Alan Mason Chesney Medical Archives of The Johns Hopkins Medical Institutions from 1978 to 1994.[9] William Jackson used citation analysis techniques to confirm that Richard Trueswell's theory (that 20 percent of the collections are used 80 percent of the time) was true at his institution.[10] Other studies focused on a defined set of users: Diane Beattie used both citation analysis and survey methods to analyze archival use of members of the Canadian Committee on Women's History (CCWH).[11] Recent citation studies, most notably the work of Donghee Sinn, have focused solely on references to digital sources.[12]

In addition, there have of course been influential archival user studies that have used entirely different methodology, including surveys and interviews of historians

---

5. Clark Elliott, "Citation Patterns and Documentation for the History of Science: Some Methodological Considerations," *American Archivist* 44, no. 2 (1981): 131–42.

6. Frederic Miller, "Use, Appraisal, and Research: A Case Study of Social History," *American Archivist* 49, no. 4 (1986): 371–92.

7. Ibid.

8. Jacqueline Goggin, "The Indirect Approach: A Study of Scholarly Users of Black and Women's Organizational Records in the Library of Congress Manuscript Division," *Midwestern Archivist* (1986): 57–67.

9. Nancy McCall and Lisa A. Mix, "Scholarly Returns: Patterns of Research in a Medical Archives," *Archivaria* 41 (1996): 156–87.

10. William J. Jackson, "The 80/20 Archives: A Study of Use and Its Implications," *Archival Issues* (1997): 133–45.

11. Diane Beattie, "An Archival User Study: Researchers in the Field of Women's History," *Archivaria* 29 (1989): 33–50.

12. Donghee Sinn, "Impact of Digital Archival Collections on Historical Research," *Journal of the Association for Information Science and Technology* 63, no. 8 (2012): 1521–37; Donghee Sinn, "The Use Context of Digital Archival Collections: Mapping with Historical Research Topics and the Content of Digital Archival Collections," *Preservation, Digital Technology & Culture* 42, no. 2 (2013): 73–86; Donghee Sinn and N. Soares, "Historians' Use of Digital Archival Collections: The Web, Historical Scholarship, and Archival Research," *Journal of the Association for Information Science and Technology* 65, no. 9 (2014): 1794–1809.

and other archival information seekers by Margaret Stieg,[13] Duff and Johnson,[14] Duff, Craig, and Cherry,[15] Helen Tibbo,[16] Margaret Stieg Dalton and Laurie Charnigo,[17] and more recently Alexandra Chassanoff.[18]

Virtually all of the citation papers mentioned (and many of the survey- or interview-based papers as well) have pointed out the limitations of citation studies. Citation studies are seen as incomplete because they cannot show the information the researcher did not find,[19] and as misleading because they don't reflect the importance of the archival material to the scholars' work.[20] The latter is especially true of reference studies like this project—which will count source materials only once, rather than each time a source was cited—because each source is assumed to have the same impact.[21] Another widely discussed apprehension to undertaking citation studies has long been summed up by the adage that researchers do not always cite what they use and do not always use what they cite.

But it is not the purpose of this study to build a full understanding of historians' information-seeking behavior in the archives; rather, the aim here is to try to provide a recent, evidence-based, wide-angle view of academic historians' use of both physical and digital archival material from a range of repositories, as such a study is currently missing from the literature.

## Methodology

I chose to look at the archival references of one of the most common groups of archival users: academic historians. Archival material is one of the sources historians consider most important,[22] which makes it probable that a good number of their books will contain a number of references to archival material. That their archival use must be revealed in citations as a necessary part of their output provides another advantage to focusing on historians over less academic archival users.

---

13.   Margaret F. Stieg, "The Information Needs of Historians," *College & Research Libraries* 42, no. 6 (1981): 549–60.

14.   Wendy Duff and Catherine Johnson, "Where Is the List with All the Names? Information-seeking Behavior of Genealogists," *American Archivist* 66, no. 1 (2003): 79–95.

15.   Wendy Duff, Barbara Craig, and Joan Cherry, "Historians' Use of Archival Sources: Promises and Pitfalls of the Digital Age," *Public Historian* 26, no. 2 (2004): 7–22.

16.   Helen Tibbo, "Primarily History in America: How US Historians Search for Primary Materials at the Dawn of the Digital Age," *American Archivist* 66, no. 1 (2003): 9–50.

17.   Margaret Stieg Dalton and Laurie Charnigo, "Historians and Their Information Sources," *College & Research Libraries* 65, no. 5 (2004): 400–25.

18.   Alexandra Chassanoff, "Historians and the Use of Primary Source Materials in the Digital Age," *American Archivist* 76, no. 2 (2013): 458–80.

19.   Jacqueline Goggin, "The Indirect Approach," *Midwestern Archivist* (1986): 57–67.

20.   Frederic Miller, "Use, Appraisal, and Research," *American Archivist* 49, no. 4 (1986): 371–92.

21.   Donghee Sinn, "Impact of Digital Archival Collections," 1521.

22.   Dalton and Charnigo, "Historians and Their Information Sources," 405.

References to archival materials in monographs have been chosen over archival references in journal articles not only because monographs are considered an important publication for historians, but because fewer citation studies have been conducted on references in books.[23] Recent use of archives is also an important factor inherent to the design of this study. For that reason, the study is limited to books published in 2012.

The 136 monographs examined were published by American university presses that have the top 50 highest total citation counts in Scopus History Journals, according to the work of Zuccala, Guns, Cornacchia and Bod.[24] The works of commercial publishers in this initial list were not included. In addition, only the 15 publishers from this list who make their full-text works accessible via Project Muse eBooks are included in analysis. Project Muse's holdings tool was used to find titles with the listed discipline of "History."[25] Books with multiple authors listed or that had the subject "Poetry, Fiction, and Creative Non-Fiction" were excluded. This left 136 books for analysis.

I took references from each book's Bibliography or Works Cited section. For those books lacking a bibliography or similar section, I looked through the notes section. In the absence of a separate notes section, I searched the full text of the book for the words "archives," "libraries," "collections," "box," "folder," and "fonds." I then captured and counted each archival repository cited, only counting a repository once per each book. Repositories mentioned in illustration credits were not counted. I also did not include microfilm collections or any publications that could be accessed in multiple places. Digital collections were only counted if they were work from an existing repository and not the archives of a publication or website.

The other data considered were the subjects associated with each book. For this I consulted the Project Muse holdings tool.[26] Their designated research areas, or "disciplines" and "subdisciplines" for each book, include time periods (such as "the Colonial Era"), places ("U.S. History>Local>South"), and subjects (for instance, "Science, Technology, and Mathematics"). Where this information was missing, or seemed to contradict the title and subtitle of the book in question, I consulted subject headings in WorldCat.

23. Blaise Cronin and Kathryn La Barre, "Mickey Mouse and Milton: Book Publishing in the Humanities," *Learned Publishing* 17, no. 2 (2004): 85–98.

24. Alesia Zuccala, Raf Guns, Roberto Cornacchia, and Rens Bod, "Can We Rank Scholarly Book Publishers? A Bibliometric Experiment with the Field of History," *Journal of the Association for Information Science and Technology* 66, no. 7 (2015): 1333–47.

25. Muse Holdings Tool, available online at muse.jhu.edu/about/librarians/holdings.html [accessed 5 October 2017].

26. Ibid.

## Findings

Ninety-three (or 68%) of the 136 books studied contained at least one reference to an archival collection. In total, there were 895 references to archival collections, and to 525 different archival repositories. The average number of repositories cited per book was 6.4, with the largest number of repositories cited in a single book being 44. Most archival collections cited were physical. Only 26 cited collections fit the criteria of a digital archival collection (digitized items housed and curated by an archival repository, and not the archives of a publication or website.)

### *Repositories Used*

I identified fifteen types of repositories, fitting within six broader categories. These are:

- Private repositories
    - ☐ Religious
    - ☐ Independent (such as Huntington Library)
    - ☐ Corporate
- Local repositories
    - ☐ Local manuscript (county, city historical societies)
    - ☐ Court
    - ☐ School (in other words, a K–12 institution)
    - ☐ City Government
- State Repositories
    - ☐ State Historical Societies
    - ☐ State Library and Archive
- National
    - ☐ National manuscript libraries (such as Library of Congress)
    - ☐ National government archives (examples: NARA, Presidential libraries)
- Non-U.S. Repositories

The last category—non-U.S. repositories—describes repositories, of all the other types, that are outside the United States. I found it necessary to create this very broad grouping for two reasons: First, I cannot claim to know enough about how repositories are arranged from country to country to properly identify them; and second, many of the citations to foreign repositories are in languages I do not read, which made discerning discrete repositories all the more difficult.

Non-U.S. repositories were used more than every category of repository except universities. Of the 122 different foreign repositories cited (as far as I could differentiate them), repositories in England, France, and Germany were most highly used. Given the same caveat (that I cannot claim to fully understand enough about

foreign repositories to disambiguate them), I discovered that citations to foreign repositories fell roughly into the same categories and in roughly the same amounts as citations to American repositories.

The highest percentage of collections cited are held at American universities and colleges: 309 of 895 (or 34%). Overwhelmingly, these collections are manuscript collections and not the institutional archives of the university itself.

Regardless of repository type, the most highly cited single institution in this sample is the Library of Congress (22 citations), followed by NARA, the University of Michigan, Yale, and the Wisconsin Historical Society.

| TABLE 1 | | |
|---|---|---|
| **Top Individual Repositories Cited** | | |
| **Repository** | **Repository Type** | **Reference Count** |
| Library of Congress | National | 23 |
| National Archives and Records Administration | National Government | 20 |
| University of Michigan | University | 19 |
| Yale University Libraries | University | 18 |
| Wisconsin Historical Society | State Historical Society | 15 |
| Stanford | University | 13 |
| UCLA | University | 11 |
| Columbia University | University | 10 |
| Harvard | University | 10 |
| UC Berkeley | University | 10 |
| NARA II | National Government | 9 |
| Duke | University | 9 |
| Huntington | Independent | 8 |
| University of Virginia | University | 8 |
| Massachusetts Historical Society | State Historical Society | 7 |
| Texas State Library and Archives Commission | State Library/Archives | 7 |
| Princeton | University | 7 |
| University of North Carolina | University | 7 |
| Franklin Roosevelt Presidential Library | Presidential | 6 |

Only 3 percent, or 26 of the 895 collections cited, were described as or appeared to be digital collections. Again, I am defining a digital collection as the output of an established repository as opposed to a database or the archives of a website or publication.

### References and Repositories Used, by Book Subject

I took into account the subject areas found in the form of subject headings Project Muse provided as well as Library of Congress subject headings assigned for some books (accessed via WorldCat). Not every title was given a subject (or in Project Muse's lexicon, a subdiscipline or research area): sometimes they were labeled simply as history books. More commonly, books were tagged with multiple subjects or subdisciplines. For instance, a book could be about the nineteenth century and about literature—and, more often than not, books with the subject of literature *were* also labeled as concerning the nineteenth century.

Subject terms or research areas can be broken down into three categories: place, time, and topic. In terms of geographical research area, 93 of the 136 books sampled concern U.S. history, while 36 are labeled as world, European, or other non-U.S. history. Forty-eight of the books are labeled as being about history of the twentieth or twenty-first century; 26 books concern the nineteenth century, and 19 deal with pre-nineteenth century history ranging from ancient to colonial times. The remaining 43 books were not categorized as concerning a specific era.

Topically, the books either concern religion (12 books), military (16), STEMM[27] (7), literature (10), or some sort of social science ranging from law to gender studies (47). The remainder of the books did not specify a subject beyond era or region.

What research areas used the most archival collections? There is not much variation across research areas. If we look at total numbers of archival citations per number of books in each subdiscipline, STEMM books used the most, but the sample size is small: there are only 7 STEMM titles. Perhaps more significantly, 84 percent of the 23 nineteenth-century titles used at least one archival collection, followed by religion, twentieth-century titles, and U.S.-specific titles.

Did different historical research areas or subdisciplines use some repository types more than others? In this admittedly very small sample, some patterns can be seen: books on military history used non-U.S. sources more than other books; religious history titles used private institutions more than other books did. Literary history books used university repositories as well as state and local repositories more often than other titles. Social science books used national repositories more than other titles.

---

27. Science, Technology, Engineering, Math, and Medicine.

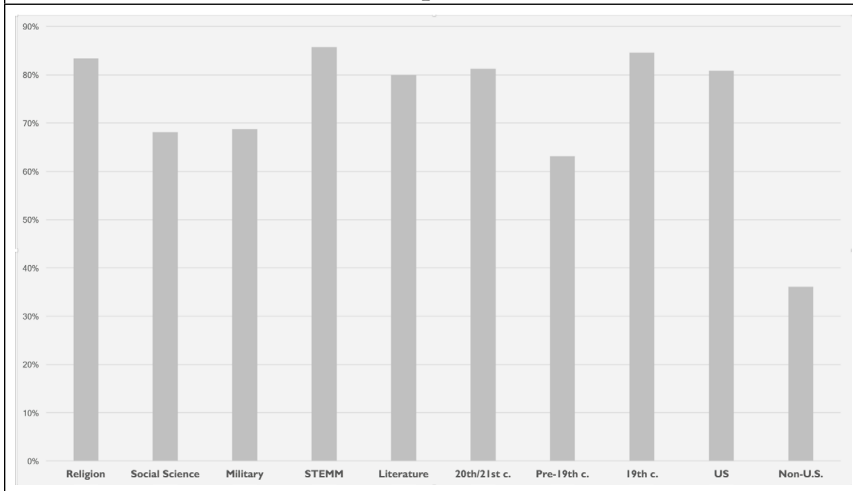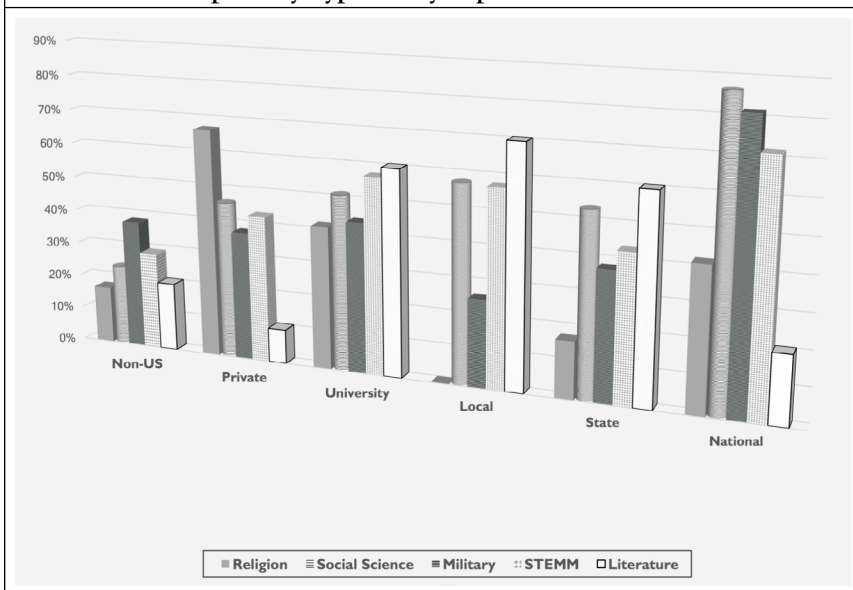**FIGURE 1**
**Total Citations per Research Area**



**FIGURE 2**
**Repository Type Use by Topical Research Area**



In terms of era, non-U.S. and private sources were cited the most in books about subjects before the nineteenth century; university, local, and state were cited the most in books about the nineteenth century; and books about the twentieth century cited the most sources from national archives. Looking at location, the only categories of books that used more non-U.S. than U.S. repositories were books that were about locations and countries outside the United States.

**FIGURE 3**
**Repository Type Use by Chronological Research Area**



**FIGURE 4**
**Repository Type by Geographical Research Area**
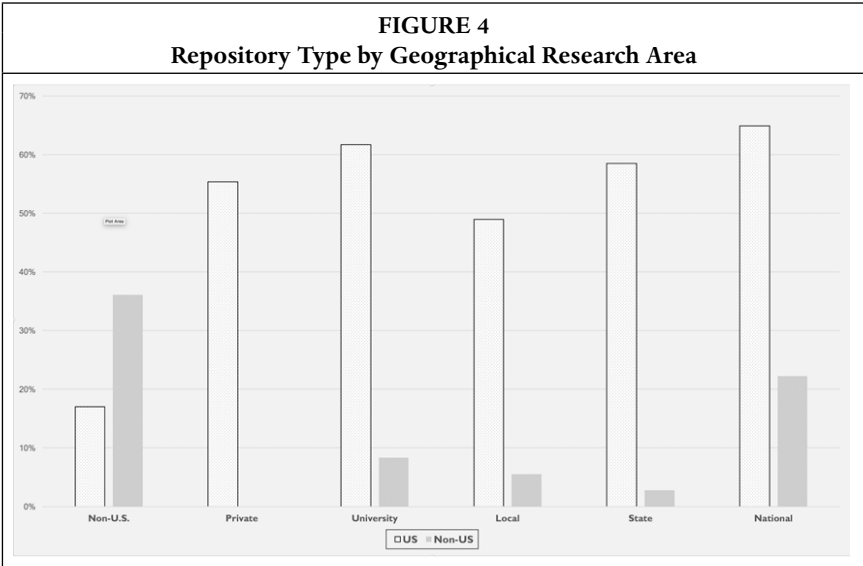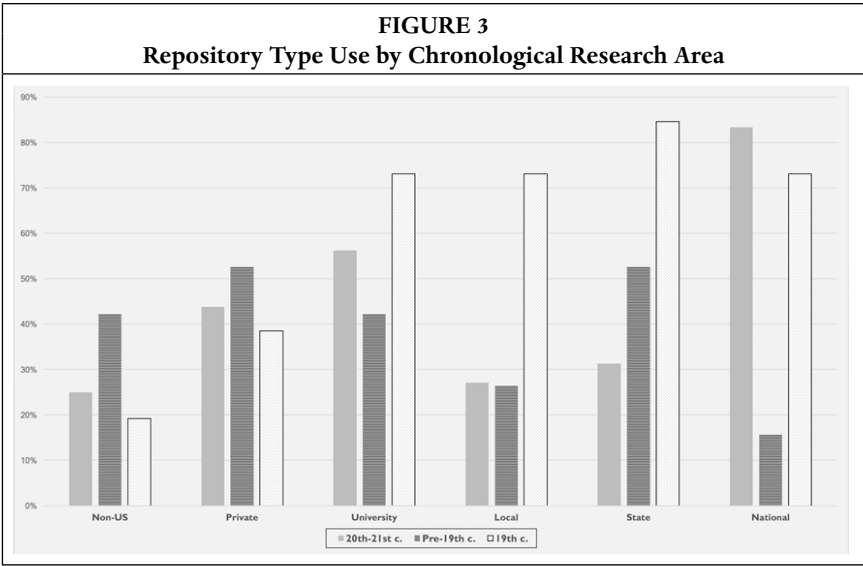
All categories of books, barring those labeled as religious histories, used university-held collections more than any other type of repository.

## Discussion

We know at least in this small sample less than three-fourths of scholarly history books cited archival collections. In terms of repository type, overwhelmingly these were citations to physical collections, and more often than not, these books cited collections from universities, non-U.S., and private institutions.

What can we tell from these numbers? The percentage of books citing archival collections is lower than the number of articles citing archival collections in Miller's 1986 study, in which 81 percent of titles cited archival collections. Considering the number of primary sources that may have been published since 1986 may account in part for the lower amount of titles using archival material in my study, but I suspect that source, or perhaps just chance, plays a larger role here. I should also point out that the percentage in my study is far higher than the percentages in Heinzkill or Elliot's studies (28% and 12%, respectively).[28]

The average number of repositories used (6.4) is almost three times higher than the average number of repositories used in Miller's study (2.2).[29] That authors use, on average, material from more institutions than in Miller's 1986 study may reflect the vastly improved ability to share repositories' holdings, largely due to finding aids and other tools that repositories have placed on the Internet.

The small number of digital collections cited, however, seems to indicate that the number of digital collections of archival materials housed and made available on the Internet has not yet reached the size or scope that would make them, on the whole, more useful than physical collections.

In terms of authors' use of individual physical repositories, most of the findings here are consistent with common sense: Numbers of citations to repositories reflect in part the structure, size, and recordkeeping missions of different archival repositories. For example, the largest and most diverse individual repositories used (such as Library of Congress) are cited the most often. A relatively large number of religious repositories are cited, which seems to reflect how church archives are kept at the diocese or parish level, as opposed as to being housed in fewer, centralized locations.

That universities were used more than other repositories reflects in part the simple fact that there are more university repositories than there are, for instance, state historical societies, private collections, or presidential libraries, and that universities may hold more diverse collections than other types of repositories such as corporate, school, or court archives.

Other possible explanations for greater use of university collections is the ease of finding information. It could easily be argued that large universities have more resources available not only to accession more materials but to make them findable to scholars.

---

28.   Miller, "Use, Appraisal, and Research: A Case Study of Social History," 379; Heinzkill, "Characteristics of References in Selected Scholarly English Literary Journals," 35.

29.   Miller, "Use, Appraisal, and Research: A Case Study of Social History," 379.

But which scholars end up at what repositories? While Miller also looked at how the citations were distributed by era and subject of articles, it is hard to compare these numbers to my study, given that in this case Miller was looking at number and types of series used (for example, personal or organization files) and not repositories used.

Looking at the results, you can see a map of where different kinds of research materials are housed: religious papers by and large are kept within their own regional institutions; material used in literary histories—which would tend toward personal or family papers—is scattered among local, state, university, and national repositories; papers of interest to STEMM, military, and social science scholars are at universities and larger national institutions. We end up with a very loose sketch of what kinds of papers are held where, as well as what kinds of information scholars in different subdisciplines are looking for.

Does this tell us anything new as archivists? The results from this small sample reveal little that is surprising. But expanding this approach into a larger study may be able to help us verify and understand patterns of use.

## Conclusions

I set out to see if looking at citations to archival collections in history books could provide a more recent and repository-agnostic picture of historians' use of archives. My investigation was able to confirm that the majority of historians (at least in this sample) do indeed use and cite archival material in their texts. I also found that historians overwhelmingly cite physical, as opposed to digital, archival collections. Predictably, we could also see from this study that amount of use of repositories by authors seems strongly correlated to the size of the repository, or rather the amount and variety of collections that the repository holds. But in terms of determining any patterns of archival use based on subdiscipline, the number of books was too small to be able to infer anything of value. (For example, STEMM history books in my sample used on average more archival repositories than did any other type of history book—but there were only 7 STEMM titles.)

Therefore, we cannot conclude from this study that archival use varies significantly across historical subdiscipline or subject. The use of vendor or library-provided subject designations to determine historical specialties in this study might have hobbled this investigation, as such designations are not designed to reflect the nuance of historical study. But it is the broad and interconnected scope of historical study itself that may make it difficult to tease out any meaningful kind of pattern concerning repositories or collections used by subject matter.

Increasing the sample size and adding questions such as how many times archives as opposed to published sources were cited could give us a more detailed baseline of archival use through references that we can use to measure against either future use or use in other disciplines. It might also be valuable to look at whether or not the low numbers of digital collections cited represents a trend across a greater body of historical titles and if we can track whether or not digital collections are seeing significant use.