

## **Searching HathiTrust: Old Concepts in a New Context**

Eamon P. Duffy

Coordinator, Humanities and Social Sciences Library

Liaison Librarian for History, Classical Studies and Government Information

McGill University

[eamon.duffy@mcgill.ca](mailto:eamon.duffy@mcgill.ca)

### ***Abstract***

This paper examines the use of the HathiTrust Digital Library as a resource for locating primary sources in the course of historical research. It focuses on the incorporation of library metadata into HathiTrust database records and the benefits this has on search strategies. The author outlines a method for locating primary sources that reduces information overload and allows researchers to target specific types of documents commonly used by historians. Historians report turning more and more frequently to open book repositories to find relevant material for their research, but also that searching full text databases often leads to information overload. The case study presented demonstrates that HathiTrust provides not only a rich collection of primary sources but also that its inclusion of Library of Congress Subject Headings creates a uniquely effective means of searching them with higher recall and precision. Despite its usefulness, the author identifies some limitations to searching LCSH in HathiTrust and makes recommendations for overcoming them. Librarians in academic environments should promote the informed use of HathiTrust among their faculty as a way of providing superior research support and improving collaboration between researchers and librarians.

### ***Keywords***

HathiTrust; primary sources; historians; Library of Congress Subject Headings; free-floating subdivisions

### ***Introduction***

Since its inception in 2008, HathiTrust has made headlines within the library world and beyond. Most of the attention focuses on the immense amount of content available within the collection and the great leap forward it represents for access to digitized print material (e.g., Christenson). The landmark October 2012 court decision it won against several authors' guilds brought significant media attention, and the outcome of the trial will have important effects on the future of access to digitized library collections ("Unlocking the Riches of HathiTrust"). Lost amid this excitement, however, is another unique aspect of HathiTrust: its transformative search capability. While other massive collections of digitized

books, such as Google Books and Open Library, also boast large keyword-searchable collections, HathiTrust is the only one to completely integrate standard library metadata into its records. This practice allows for more robust content searching and a way to mitigate the problems inherent in keyword searching large bodies of text.

Librarians have tremendous experience in combining keyword searching with controlled vocabulary in other contexts, such as full-text article databases. The opportunity to use this strategy when searching books is revolutionary. This paper focuses on the inclusion of Library of Congress Subject Headings (LCSH) in HathiTrust records, with particular emphasis on free-floating subdivisions. It will demonstrate their use as a means of overcoming information overload, reducing the number of irrelevant hits, and limiting searches to specific types of documents, while also pointing out some limitations of controlled vocabulary within HathiTrust.

Perhaps the greatest beneficiaries of the HathiTrust Digital Library's search capabilities are historians. They rely mainly on primary sources, which are documents written in the time period being studied. These documents offer first-hand accounts of events and provide the most reliable information about the past. HathiTrust is a natural choice for this research owing to its historical depth. Historians, in particular, can utilize its search features to more easily identify relevant primary sources, as well as to discover "hidden" ones.

### ***About HathiTrust***

HathiTrust is a massive open digital repository of books scanned from the collections of nearly 70 research libraries across North America and in Europe. It takes its name from the Hindi word for "elephant," an animal celebrated for its long memory and thus symbolizing the repository's commitment to preserving digital material. The current collection contains more than 10.6 million volumes (HathiTrust, "Currently Digitized"). The original core of the collection consisted of the scanned images created from some member institutions' collections during the Google Books Library Project. It has now grown to include content digitized through projects involving the Internet Archive and Microsoft's Live Search Books program as well as through local digitization efforts by HathiTrust members.

HathiTrust's goal is two-fold: the long-term preservation of content digitized from members' print collections and improved access to these documents (HathiTrust, "Mission and Goals"). About thirty percent of titles in HathiTrust are in the public domain. Anyone can freely read these documents online and download PDFs of individual pages. Researchers from HathiTrust member institutions have the added option of downloading PDFs of entire public domain documents. The remainder of titles in HathiTrust fall under copyright restrictions. They are full-text searchable by the public at large, though content is only available through limited views.

Collaboration among libraries is an important aspect of HathiTrust. Members in the partnership have a voice in shaping its development and in making decisions about what material from their local collections will be added to the digital library. Because it was created by libraries, and aims to assist its members with content preservation and the provision of access to their users, HathiTrust's metadata structure includes information not systematically incorporated into other open digital book repositories. All items uploaded to HathiTrust must be accompanied by a MARC record (HathiTrust, "Guidelines for Digital Object Deposit"). One primary purpose of this requirement is to allow libraries to easily integrate HathiTrust content into their local library catalogues. Fortunately, it also has secondary benefits to HathiTrust end-users.

The marriage of LCSH and full text – not available in similar open repositories – allows for incredibly powerful and efficient search strategies. It reduces the information overload and irrelevant results normally associated with full text searching of such a large corpus while simultaneously providing access to relevant material normally missed when restricting searches to library catalogues. These new search opportunities could have a transformational impact on the quality of search results and the information searching habits of historians and other researchers.

### ***Historians' information behaviour***

Studies have consistently shown that historians profess to rely most heavily on archives, special collections and other sources of unpublished documents to complete their research (e.g., Rutner and Schonfeld 8), yet the results of quantitative citation analyses obscure that assertion. When Stieg Dalton and Charnigo surveyed 278 historians about their information behaviour, they found that although respondents reported that archival material and unpublished manuscripts constituted the types of material they used most often as primary sources, citation analysis of historical writing published in 2001 showed that published material like books and newspaper articles made up the majority of cited works within historical writing (405). Another more in-depth citation analysis of articles appearing in the *American Historical Review* between 2001 and 2010 showed that published material (representing both primary and secondary sources) vastly outnumbered archival and special collections material within article citations (Sinn 1526). It is clear, then, that published works play an important role as primary sources in historians' research, and tools that better facilitate their discovery would benefit scholarship.

Though they are among the most intense users of academic library collections (Delgadillo and Lynch 245), as well as being among their most passionate supporters, several studies conducted over the past thirty years report that historians rarely interact with librarians when conducting research (Stieg 554; Stieg Dalton and Charnigo 408-09; Rutner and Schonfeld 20-22). The reasons historians give for their lack of consultation vary, though librarians' lack of subject knowledge appears most often among them. Perhaps not surprisingly to

librarians, these same studies sometimes also highlight the fact that historians undertake their research unsystematically and without using resources that could greatly facilitate their work. This ironic state of affairs – where both parties operate independently due in part to a perceived lack of knowledge, despite the fact they have so much to learn from other – contributes to a climate of genuine but detached respect. The distant working relationship between librarians and historians results in scholarship that might arguably have been better and library service that could certainly have been more effective. Even in cases where they lack detailed subject knowledge, history librarians need to find ways to share their knowledge of information sources and how to use them effectively.

One important new development is that historians are now embracing the electronic environment. Even ten years ago historians reported preferring electronic resources to locate primary and secondary sources (Stieg Dalton and Charnigo 412). A major 2012 study surveyed thirty-nine academic historians and graduate students about their research habits and experiences. Though the report made no direct mention of HathiTrust, the authors did include a section on historians' use and impression of another major digitized book resource, Google Books (Rutner and Schonfeld 18-20). Nearly all historians interviewed mentioned using Google Books in their research, especially those studying periods prior to 1923, since the full text of works published before this date are available because they no longer fall under US copyright restrictions. They also mentioned the transformational effect that massive book digitization has on their work. As one interviewee put it, "Being able to search for a particular word that I'm interested in is so much more powerful than searching a library catalog. It's not in any title. It's not in a subject term" (19).

Though convenient, these large collections do pose some practical difficulties for many historians. In a 2010 paper, Allen and Sieczkiewicz looked at the ways in which historians interact with a specific type of primary source: digitized historical newspapers. Though the historians they interviewed found these resources could be a convenient alternative to microfilmed copies, they also identified several limitations of keyword searching large quantities of text as a research strategy (3). The first of these was information overload. Due to the sheer amount of searchable text, searches often retrieve too many results for the researcher to analyze completely. Another limitation historians mentioned was their inability in some cases to limit searches to specific types of articles. Whether a searched term appeared in an advertisement or an editorial may be an important distinction, and ideally databases would allow researchers to limit their searches in this way. Due in part to these limitations, one interviewee described historians "as people 'who would rather browse' an entire year of newspapers on film than search for specific articles" (3). Stieg Dalton and Charnigo identified similar difficulties historians have with online databases. Of the types of problems they reported to encounter, 44 percent related to either information overload or unsatisfactory indexing and terminology (412).

The literature on historians' search behaviour shows that while they rely heavily on books as primary sources for their research, and have quickly embraced digitized book repositories as research tools, they often complain about being overwhelmed with the volume of results they get and the poor indexing features of full text databases. Despite this, few historians report turning to librarians for assistance.

### ***Locating Primary Sources***

Primary sources are documents written during the time period a historian is studying. From a historian's perspective, their proximity to the period of events makes them more reliable than secondary sources, which interpret the past after the fact. They can be used to trace specific events or to provide clues about the social, cultural and political atmospheres at a given period in history. Serious research cannot be conducted without them. They are the cornerstone of the historical method and indispensable for supporting any argument. Primary sources come in two major varieties: unpublished material (archival records or special collections material of which one or few copies exist) and published works (books, newspaper articles, etc.). While just about any document can potentially be a primary source, there are several commonly used types. These include travel narratives, diaries, memoirs, letters, and government documents.

Finding primary sources for one's topic presents a variety of challenges and traditionally requires a mixture of persistence, patience and luck on the part of the searcher. Perhaps the simplest method is to rely on the work of one's predecessors, and identify primary sources through citations in previous works written on a topic. Following citation trails is a method historians frequently use to locate material (Rutner and Schonfeld 17; Stieg Dalton and Charnigo 408; Stieg 554). While fairly straightforward, scholars using this method alone risk derivativeness and may miss the opportunity to unearth previously unknown documents that could advance scholarship in their area of study. In the case of very novel scholarship and emerging fields, this technique might not even be an option.

Another method of finding primary sources is pure chance. Serendipity is a surprisingly common way for historians to find useful documents and therefore should not be underestimated as a "strategy." In several studies on the research habits of academic historians, serendipitous discovery of sources through browsing library stacks, for example, appears repeatedly as a method historians employ (Quan-Haase and Martin 456-57). However common this might be, historians should not rely solely on chance in their quest for primary documents. Inevitably, they must begin searching for them in earnest.

It is at this point that searchers can run into difficulty using traditional search tools – especially if there are time constraints. Searching for primary sources requires patience and involves search techniques that differ from those used to locate secondary sources. Typically, one searches for books on a topic using terms that

describe what those books should be about. When searching for primary sources, however, the trick is to identify works in which the historian would likely find useful or valuable information to interpret past events, regardless of what those works are actually “about.” This partly explains why historians mentioned the need for the ability to search for specific types of newspaper articles, rather than just by keyword, in the Allen and Sieczkiewicz study.

Librarians traditionally suggest using LCSH when searching for primary sources in a library catalogue. Paradoxically, however, the main headings are somewhat less important in this situation. Rather, the key is to use certain free-floating subdivisions that follow the main heading and identify several *document types* that often prove valuable as primary sources, such as:

- Biography (includes autobiographies and memoirs)
- Correspondence (for collections of letters)
- Description and travel (for travel narratives)
- Diaries
- Interviews
- Personal narratives
- Sources (for published anthologies of diverse primary sources on a topic)

This method of searching, though perhaps largely unknown and underused by historians, is a mainstay of library literature on the topic (Presnell 101; Kitchens 42-43). Furthermore, in his paper on the inclusion of LCSH in *Eighteenth Century Collections Online*, Jeffrey Garrett highlighted the ability to search for words contained within LCSH, and not just the main heading itself, as a major asset (71).

### **Case Study**

The following example will demonstrate the richness of the HathiTrust collection, its potential usefulness to historians, and the value of its unique search capabilities. Suppose a historian is researching the Peruvian guano industry in the nineteenth century, which supplied the world with much of its agricultural fertilizer. He wants particularly to locate primary sources describing the working conditions in Peruvian guano mines. Unless an author alive at the time happened to write a book on that specific topic, a keyword search in the library catalogue would likely turn up few, if any, relevant documents. This is because, apart from the title, LCSH provide the other major source of searchable topical text. The Library of Congress' *Subject Headings Manual*, which guides cataloguers in applying LCSH to records, states that they should only apply subject headings to topics “that best summarize the overall contents of the work and provide access to its most important topics” (H180 sec. 1). In practice, they are assigned to “topics that comprise at least 20% of a work” (H180, sec. 1). Therefore, even if the search is successful, it will miss books not wholly concerned with the topic, but which nevertheless contain brief accounts or even large sections covering it.

One such example would be the account of a traveler to Peru who might have visited a mine at one point along his journey. Travel narratives are an important primary source for several reasons. They are sometimes the best source of first-hand information about societies with small populations and low literacy levels. For researchers without foreign language skills – which is often the case with undergraduate students – travel narratives in one’s native language might be the only accessible primary source covering foreign events. Finally, these documents can provide an alternate or outsider’s perspective on a specific time and place.

As previously mentioned, the traditional method of locating travel narratives in a library catalogue is to use the Library of Congress free-floating subdivision that cataloguers use to designate these books (“description and travel”) along with the name of the place visited (Presnell 52). Using this technique, a subject search for “Peru – description and travel” will retrieve a list of travel narratives about that country. The search could be limited by date to ensure the narrative was published during or soon after the time period being studied. In order to get a comprehensive list of relevant sources, the historian would consult each found title to see if there are any sections that deal with workers in the guano mining industry. Depending on the number of search results, this verification could present a substantial obstacle. Those books without good indexes would have to be read, or at least skimmed, in their entirety. Obviously, this requires a daunting amount of effort, not to mention a significant time investment, with no promise of success. Lacking patience, or help from research assistants, the large quantity of material acts as a significant deterrent to the use of these potentially valuable resources. As a practical matter, historians conducting research in this fashion would likely target a few promising titles, leaving other potentially useful sources untouched.

With the advent of large digitized book repositories, researchers now have other means of locating books beyond the traditional library catalogue. Since they allow full-text searching of their entire contents, one might expect that this enhanced capability would allow researchers to find useful primary sources that may otherwise have remained hidden. By breaking down the case study topic into concepts, the following might be a useful string of keywords to search in all fields: peru\* guano mine\* travel\*. But even with a topic as seemingly narrow as the one chosen for this case study, searching such a large amount of text (10.6 million volumes worth in HathiTrust) might create an overwhelmingly large results list with low precision. Experiences like this lead to historians’ complaints about information overload and the frustrations of full text searching. Faced with such a mountain of results, those who do not give up might choose the first few titles that seem to meet their needs before moving on.

HathiTrust, however, permits a third search strategy – a combination of keywords and LCSH that retains the high recall of the former and produces higher precision than the latter. This can be done using the “Advanced Full-text Search.” Figure 1 demonstrates how a topical keyword search can effectively be combined with an

LC term and the “description and travel” free-floating subdivision, and limited by date.

**Advanced Full-text Search:**  
 Search information *within or about* an item  
[Search Tips](#)

in   
   in

[+ Add a pair of search fields](#)

**Limit to:**

Full view only    Year of publication:   and

Figure 1. Combined LC subject and full-text keyword search, limited by date

This search has the effect of limiting a keyword search’s results set to only those documents representing a specific document type – in this case travel narratives. Compared to a subject search or keyword search alone, the results will be much more manageable. To demonstrate this, the author completed each of the three searches described above in the HathiTrust Digital Library. The results in Table 1 are revealing.

Table 1. Results of searches in HathiTrust for travel narratives published between 1800 and 1914 that mention guano mines in Peru

Subject only	Keyword only	Combined search
362	46,010	150

A subject search alone retrieves over 360 books, which would be an incredible number to have to scan through looking for mentions of guano mines. Perhaps even more surprising, though, is the astonishing 46,010 results found via keyword. This number is a testament to the massive amount of material available in HathiTrust and its value as a repository of nineteenth century books. But it also highlights the frustrations that can arise when relying on keywords for even such a specialized topic as this one. Relevant travel narratives are lost in a sea of books on such related but irrelevant topics as the use of Peruvian guano in agricultural applications and government documents on territorial disputes concerning islands with lucrative guano deposits. The combined search, on the other hand, returns a much less intimidating set of 150 titles. Not only is this a more manageable number of results to work with, but the researcher is also sure

that each one is a travel narrative of Peru from the nineteenth century that makes at least one mention of guano or guano mining.

HathiTrust even provides researchers with the ability to quantitatively determine which of these 150 results are likely to be the most useful to one's research. Each title is individually searchable, with results indicating how many times a word or phrase appears in the book and on which pages. The more often the word "guano" is mentioned in a book, the more likely it would be worth investigating. Figure 2 shows the results of a successful search within a promising primary source whose non-descript title does not clearly advertise its relevance to the topic.

**About this Book**

Melbourne, and the Chincha Islands; with sketches ... Peck, George Washington, 1817-1859.

[View full catalog record](#)

Copyright: [Public Domain, Google-digitized.](#)

**Get this Book**

[Find in a library](#)

[Download whole book \(PDF\)](#)

Partner login required

**Add to Collection**

Login to make your personal collections permanent

Select Collection

Add

**Share**

Permanent link to this book

<http://hdl.handle.net/2027/hvd.321>

Search in this text

**guano** matched 31 pages in this item.

Viewing results for: 1 to 10 of 31 pages

1 | [11](#) | [31](#) | [next](#) ▶

**p.209** - 4 matching terms

...m the **guano** is brought by wheelbarrows and dirt cars, drawn by donkeys, from the sides of the hill where it is cut away. Coolies, who are obliged to wear thick bandages over their mouths, push the **guano** down to the lower ends of the mangueras, where there are openings connected with "shutes," or lon...

...g was **guano**, but we smoked paper cigarettees, and said " Que cuidar?" and made our- selves intelligible to each other in broken Spanish mended with worse Latin. I was satisfied with a very cursory examination of the internal economy of Senor Serrates' kingdom. Towards the cliff, along the lee side o...

...e the **guano** is conducted into launches, or directly into the holds of vessels loading. When thus taken in by launches, it has to be hoisted in over the ship's side in sacks or tubs, by a tackle and fall, a work requiring several weeks to com- plete a cargo. When poured directly into ships, moored fo...

**p.167** - 3 matching terms

...n the **guano**, about a foot beneath the surface, and three feet within. The **guano** is everywhere perforated and honey- combed with these holes, and the birds may be taken from them any time by reaching in the hand. One day when we went gunning, two men took out more than fifty in a short time. I carrie...

Figure 2. Searching within a title

This feature is available whether or not the book is under copyright, so the value of searching HathiTrust for primary sources is not limited to those available in full text. In fact, historians interviewed by Rutner and Schonfeld reported that they already use the information available in limited and snippet views within Google Books in a similar way (19). Just knowing where a particular word or phrase appears within a book, and how often, helped them determine whether locating a print version of that title would be worthwhile.

## Limitations

Controlled vocabularies inherently contain certain limitations. There are many criticisms of searching LCSH in particular, and these have been covered in depth in three major review articles covering literature back to 1944 (Kirtland and

Cochrane; Shubert; Fischer). Some, such as the biases evident in the wording of LCSH, are significant but outside the scope of this article. Others, such as the fact that cataloguing practice evolves over time, must not be forgotten. New headings and subdivisions are created, and the rules governing their use change as well. For example, *description and travel* now encompasses all travel writing, while before 1991 works describing cities or colonies were assigned the now defunct free-floating subdivision, *description*. Additionally, while the *personal narratives* subdivision was used to describe autobiographical material in a number of contexts prior to 1977, since that year the *Subject Headings Manual* restricts its usage to events and wars. The free-floating subdivision currently used with classes of persons, types of activities and diseases is *biography*. Ideally, older catalogue records would be updated to incorporate these changes, but in reality this rarely occurs. Many HathiTrust MARC records contain elements dating from the time the original print version was catalogued. Therefore, searchers must be mindful of this fact and try multiple terms in some situations.

In order to reap the advantages of the search methods described in this paper, the ability to cross-search keywords with subject headings is not enough on its own. Search results are only as good as the information included in the MARC records that accompany each HathiTrust title. At least one author has already noted a general trend towards vague and general subject headings within catalogue records and how this practice “can frustrate historians, their students, and the reference librarians trying to assist them” (Kitchens 43). Indeed, the *Subject Headings Manual* does instruct cataloguers to “assign headings that are as specific as the topics they cover” (H180 sec. 4). Without local cataloguing practices that encourage specificity in assigning subject headings and the use of elements such as free-floating subdivisions, however, the method described in this study will simply not work.

Furthermore, while HathiTrust requires each uploaded title to have an accompanying MARC record, and its interface provides a means of subject searching, subject headings are neither required nor “strongly preferred” elements of acceptable records (HathiTrust, “Bibliographic Metadata Specifications”). To verify just what percentage of titles in HathiTrust contain LCSH in their records, the author analyzed the MARC records for all public domain books available through the HathiTrust OAI feed from the University of Michigan as of December 16, 2012. For the purposes of this examination, any record containing at least one instance of a 600, 610, 611, 630, 650, 651 or 655 field with a second indicator value of “0” was deemed as having a subject headings field containing LCSH. Out of a total of 1,613,486 records, roughly three-quarters (1,181,764) contained LCSH. This indicates that a significant percentage of HathiTrust records have no subject headings at all and cannot be located using a combined subject and keyword search.

## **Recommendations**

In order to preserve the integrity of the enhanced search access that makes HathiTrust unique among its peers, and so potentially valuable to researchers, member libraries should be encouraged to include subject headings in records uploaded to HathiTrust. Perhaps the obvious benefits of subject headings within HathiTrust will provide some motivation for their inclusion in its records. Furthermore, HathiTrust should do more to promote the fact that it uses LCSH as a controlled subject vocabulary since this is not immediately evident to the average researcher. One method would be the option to browse the LCSH thesaurus as one can in traditional OPACs. As it stands now, one needs to have quite sophisticated familiarity with LCSH in order to incorporate it effectively into searches.

Historians cannot benefit in the ways outlined in this paper if they do not learn how to do so. It is not likely that many will ever do this on their own, and research shows that historians do not make a habit of consulting librarians for research advice. History librarians are encouraged to share the information presented here with their faculty members as one step towards fostering more collaborative relationships. This can be accomplished in several ways, such as adding information about free-floating subdivision searching within HathiTrust into history subject and course guides. Some history departments hold local informal meetings or workshops where faculty and graduate students can discuss important issues in the field. Such a session on digital humanities or the use of technology in historical research would provide an ideal venue for a short presentation on searching HathiTrust. Librarians' search skills and knowledge of resources complement historians' deep subject familiarity, which together can offer new opportunities to scholars of history.

Finally, this paper only scratches the surface of the transformative effect that HathiTrust can have on academic research. Beyond the field of history, other humanists and social scientists will also find similar value in the ability to search text formerly hidden within book covers. Librarians need to discover and publicize how HathiTrust and other digitized book repositories can benefit researchers in their specific areas.

## **Works Cited**

- Allen, Robert B., and Robert Sieczkiewicz. "How Historians Use Historical Newspapers." *Proceedings of the American Society for Information Science and Technology* 47.1 (2010): 1-4. Web. 25 April 2013.
- Christenson, Heather. "Hathitrust: A Research Library at Web Scale." *Library Resources & Technical Services* 55.2 (2011): 93-102. Web. 25 April 2013. <<http://www.hathitrust.org/documents/christenson-lrts-201104.pdf>>.

- Delgadillo, Roberto, and Beverly P. Lynch. "Future Historians: Their Quest for Information." *College & Research Libraries* 60.3 (1999): 245-59. Web. 25 April 2013. <<http://crl.acrl.org/content/60/3/245.full.pdf+html>>.
- Fischer, Karen S. "Critical Views of LCSH, 1990–2001: The Third Bibliographic Essay." *Cataloging & Classification Quarterly* 41.1 (2005): 63-109. Web. 29 April 2013.
- Garrett, Jeffrey. "Subject Headings in Full-Text Environments: The ECCO Experiment." *College & Research Libraries* 68.1 (2007): 69-81. Web. 25 April 2013. <<http://crl.acrl.org/content/68/1/69.full.pdf+html>>.
- HathiTrust. "Bibliographic Metadata Specifications." HathiTrust, n.d. Web. 25 April 2013. <[http://www.hathitrust.org/bib\\_specifications](http://www.hathitrust.org/bib_specifications)>.
- . "Currently Digitized." HathiTrust, n.d. Web. 25 April 2013. <<http://www.hathitrust.org/about>>.
- . "Guidelines for Digital Object Deposit." HathiTrust, n.d. Web. 25 April 2013. <[http://www.hathitrust.org/deposit\\_guidelines](http://www.hathitrust.org/deposit_guidelines)>.
- . "Mission and Goals." HathiTrust, n.d. Web. 25 April 2013. <[http://www.hathitrust.org/mission\\_goals](http://www.hathitrust.org/mission_goals)>.
- Kirtland, Monika and Pauline Cochrane. "Critical Views of LCSH - Library of Congress Subject Headings, A Bibliographic and Bibliometric Essay." *Cataloging & Classification Quarterly* 1.2/3 (1982): 71-94. Web. 29 April 2013.
- Kitchens, Joel D. *Librarians, Historians, and New Opportunities for Discourse : A Guide for Clio's Helpers*. Santa Barbara, CA: Libraries Unlimited, 2012. Print.
- Library of Congress. Cataloging Policy and Support Office. "H 180: Assigning and Constructing Subject Headings." *Subject Headings Manual*. Washington: Library of Congress, 2008-. Cataloger's Desktop. Web. 29 April 2013.
- Presnell, Jenny L. *The Information-Literate Historian: A Guide to Research for History Students*. New York: Oxford University Press, 2007. Print.
- Quan-Haase, Anabel, and Kim Martin. "Digital Humanities: The Continuing Role of Serendipity in Historical Research." *Proceedings of the 2012 iConference*. 7-10 Feb. 2012, Toronto. New York: ACM, 2012. Web. 25 April 2013.
- Rutner, Jennifer , and Roger C. Schonfeld. *Supporting the Changing Research Practices of Historians*. N.p. : Ithaka S+R, 2012. Web. 25 April 2013.

<http://www.sr.ithaka.org/research-publications/supporting-changing-research-practices-historians>>.

Shubert, Steven Blake. "Critical Views of LCSH - Ten Years Later: A Bibliographic Essay." *Cataloging & Classification Quarterly* 15.2 (1992): 37-97. Web. 29 April 2013.

Sinn, Donghee. "Impact of Digital Archival Collections on Historical Research." *Journal of the American Society for Information Science and Technology* 63.8 (2012): 1521-37. Web. 25 April 2013.

Stieg Dalton, Margaret, and Laurie Charnigo. "Historians and Their Information Sources." *College & Research Libraries* 65.5 (2004): 400-25. Web. 25 April 2013. <<http://crl.acrl.org/content/65/5/400.full.pdf+html>>.

Stieg, M. F. "The Information of Needs of Historians." *College & Research Libraries* 42.6 (1981): 549-60. Print.

"Unlocking the Riches of HathiTrust." *American Libraries* Jan./Feb. 2013: 40-43. Web. 25 April 2013. <<http://americanlibrariesmagazine.org/features/01152013/unlocking-riches-hathitrust>>.