

Fall 2015

## Identity and User Preference in the Presentation and Content of Digital Archives: A Study of the Plumas County Museum's Haun Collection

Lindsay Morton  
*San Jose State University*

Follow this and additional works at: [https://scholarworks.sjsu.edu/etd\\_theses](https://scholarworks.sjsu.edu/etd_theses)

---

### Recommended Citation

Morton, Lindsay, "Identity and User Preference in the Presentation and Content of Digital Archives: A Study of the Plumas County Museum's Haun Collection" (2015). *Master's Theses*. 4654.  
DOI: <https://doi.org/10.31979/etd.h93a-g5zx>  
[https://scholarworks.sjsu.edu/etd\\_theses/4654](https://scholarworks.sjsu.edu/etd_theses/4654)

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact [scholarworks@sjsu.edu](mailto:scholarworks@sjsu.edu).

IDENTITY AND USER PREFERENCE IN THE PRESENTATION AND CONTENT  
OF DIGITAL ARCHIVES:  
A STUDY OF THE PLUMAS COUNTY MUSEUM'S HAUN COLLECTION

A Thesis Presented to  
The Faculty of the School of Information  
San José State University

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Library and Information Science

by

Lindsay Morton

December 2015

© 2015

Lindsay Morton

**ALL RIGHTS RESERVED**

The Designated Thesis Committee Approved the Thesis Titled

IDENTITY AND USER PREFERENCE IN THE PRESENTATION AND CONTENT  
OF DIGITAL ARCHIVES:  
A STUDY OF THE PLUMAS COUNTY MUSEUM'S HAUN COLLECTION

by

Lindsay Morton

APPROVED FOR THE SCHOOL OF INFORMATION

SAN JOSÉ STATE UNIVERSITY

December 2015

Patricia Franks	School of Information
Debra Hansen	School of Information
Henry Lowood	School of Information
Alyce L. Scott	School of Information

## ABSTRACT

### IDENTITY AND USER PREFERENCE IN THE PRESENTATION AND CONTENT OF DIGITAL ARCHIVES:

#### A STUDY OF THE PLUMAS COUNTY MUSEUM'S HAUN COLLECTION

by Lindsay Morton

This thesis explores user preference in the presentation and content of online archives in small, local institutions. To obtain data for this study, a collection from the Plumas County Museum in northern California was digitized, and three versions were presented on a custom-built test website: 1) a straightforward reproduction of documents in the collection; 2) a pairing of reproductions and typed transcripts; and 3) a selective, interpretive reproduction with supporting secondary material. Users with a variety of research backgrounds viewed the website and provided feedback through an anonymous, online survey. Google Analytics was also used to measure site traffic.

During the five-week testing period, 25 complete surveys, five partial surveys, and traffic information from 183 unique users were gathered. Survey findings indicate that 46 percent of users found version 3—the highly processed, highly contextualized presentation—most useful. When controlling for research experience, scholarly and professional users preferred the straightforward reproduction (version 1), while students and teachers preferred an enhanced presentation (versions 2 and 3). Avocational researchers did not show a clear preference. Site traffic showed a heavy concentration of users (68%) from California, as well as users from 15 other states. These findings suggest that while local archives may be most relevant within their geographical range, digitization of collections can extend an archives' usership more broadly.

## ACKNOWLEDGEMENTS

Many people contributed to the development and execution of this study. I would like to begin by thanking the members of my thesis committee for their insight and oversight: Patricia Franks for guiding me through the process from beginning to end; Debra Hansen, for her thoughtful and constructive recommendations; Henry Lowood, for his comments on the underlying conceptual framework of the study; and Alyce Scott for her feedback on digitization processes.

I am grateful for the constant warm welcome, guidance, and support of the staff at the Plumas County Museum, Executive Director Scott Lawson and Museum Assistant Carol Paoli. I also wish to thank the Plumas County Museum Association for allowing me to reproduce artifacts from the museum's holdings for the purposes of this study.

I would like to thank my friends and family for their encouragement, flexibility and understanding, with special thanks to my mother Tamara Wilton for her tireless support and for her help in pre-screening the website before its release.

Finally, I am enormously indebted to the researchers and members of the public who took time from their busy schedules to participate in this study. This thesis would never have been possible without their enthusiasm and generous collaborative spirit.

## Table of Contents

List of Tables .....	viii
List of Figures .....	ix
Chapter 1: Introduction .....	1
The Limits of Information on Small Archives .....	3
Barriers to Digitization at Small Archives .....	7
The Move Toward Digitization .....	8
The Research.....	14
Thesis Organization .....	15
Chapter 2: Literature Review .....	16
Emerging Norms .....	16
Challenges of Theory and Practice in the Field .....	19
Archival arrangement and representation .....	19
Processing standards and practices .....	21
The question of usership .....	24
Impact of digitization on research quality .....	27
Conclusions from the Literature .....	29
Chapter 3: Study Design and Methodology .....	33
Research Setting .....	33
The Haun Collection and Reasons for its Selection for this Study .....	34
Test Environment .....	38
Platform .....	42
Site layout .....	44

Digitization and content preparation .....	45
Outreach to Potential User Groups .....	48
Data Collection .....	50
Limitations and Delimitations .....	53
Chapter 4: Results and Discussion .....	55
Chapter 5: Conclusion .....	68
References .....	74
Appendix A: Haun Collection Test Site Assessment .....	83
Appendix B: Survey Results .....	86



## List of Tables

Table 1. WordPress plugins .....	43
Table 2. User group representation .....	55
Table 3. User perceptions of overall usefulness .....	56
Table 4. Likelihood of visiting similar sites in future .....	56
Table 5. Comparison between assessment of usefulness and interest in similar sites by user group .....	57
Table 6. Usefulness assessment by user group .....	58
Table 7. Version preference .....	58
Table 8. Version preference by user group .....	59
Table 9. Frequency of consulting primary source material (per year) .....	59
Table 10. Negative written feedback by category .....	60
Table 11. Positive written feedback by category .....	60
Table 12. Site access by version .....	64
Table 13. Most accessed site pages in order of page views .....	66

## List of Figures

Figure 1. An example of the version 1 presentation with inset. view of individual record. ....	39
Figure 2. An example of the version 2 presentation. ....	40
Figure 3. An example of the version 3 presentation with sample notes feature. ....	42
Figure 4. A diagram of site layout. ....	44
Figure 5. An example of the photography set up. ....	45
Figure 6. Examples of social media outreach via Facebook (left) and Twitter (right) ....	50
Figure 7. Sample survey screen. ....	52
Figure 8. Site visits over time. ....	63

## **Chapter 1: Introduction**

Over the past 25 years, archives and special collections have increasingly embraced digitization. Early efforts focused on producing digital finding aids for analog collections (Cruikshank, Daniels, Meissner, Nelson, & Shelstad, 2005). Finding aids facilitate the discovery of information within a collection of records. Examples include card indexes, calendars, guides, inventories, shelf and container lists, and registers (Society of American Archivists, 2015). Projects that are more recent have expanded the scope of digitization to include curated online exhibitions, mass digitization of complete collections, and even large-scale collaborative databases containing resources from multiple institutions.

The benefits of archival digitization are well known. Digitization complements archival values and goals by advancing the core archival principles of preservation and access. Digital surrogates protect records from the wear and tear of regular consultation and serve as a backup in case of disasters such as fire and flooding (Gracy & Kahn, 2011). Digitized collections also have the potential to extend access to users beyond the institution. Indeed, administrators, donors, and researchers have come to expect digitization as a standard component of the archival process (Zastrow, 2013). Patrons likewise have grown to depend on the Internet, both for identifying and locating resources and for accessing material remotely. “A new generation of historical researchers—tech-savvy and accustomed to immediate access to electronic sources—is driving an expanding need for the creation of digital archives” (Hecht, 2014, p. 36). “Not only is there a belief that the public is entitled to access resources, but that increasing

awareness of these sources can be mobilized as a marketing tool” (Roth-Katz, 2012, p. 123-124).

The Internet has become a space especially suited to subcultures, niche topic areas, and specialized communities where diffuse populations with specific interests can forge connections, exchange ideas, and access resources relevant to them. “Without geographical limitations, even the narrowest interest groups can obtain enough of a critical mass to make it [forming a group] worthwhile” (Leboff, 2011, p. 123). “Online communities can be especially useful for niche topics where community members have specific needs or require specialized interactions” (Bik & Goldstein, 2013, p. 1).

In archives however, digitization has not been widely embraced as a strategy for maximizing the potential of niche audiences. Larger organizations with collections of broad historical interest are more likely to offer web-based archives than their smaller counterparts (Institute of Museum and Library Services, 2006). Many factors contribute to this difference. Smaller organizations have limited budgets, fewer staff, and a lower ratio of professional to paraprofessional personnel to undertake digitization projects. In addition, organizations designed to serve the needs of a specific community or geographic region may doubt the value of providing digitized records for users who presumably already have access to the physical collection.

Archivists have developed—or not developed—digital finding aids, exhibitions, and collections “largely based on the perceptions that archival professionals have about user needs” rather than systematic research into what archive users actually require (Allison-Burnell, Yakel, & Hauck, 2011, p. 68). Digitization remains “on the whole, an activity which only pays partial attention to the users for whom the content is being

digitized, and *their relevant needs*” (emphasis in the original) (Marchionni, 2009, para. 2). There is still much to learn about usage patterns for digital archival content, especially content with a narrow focus or specific regional significance. Decisions about digitization and the online presentation of digital archival content have enormous implications for access and may determine the long-term relevance of historical institutions. For organizations lacking the funding to support mass digitization or staff time to devote to grant seeking, it is essential that digitization programs be rooted in a sound understanding of the institution’s own particular users. The current study explores user preference in online archival presentation using both quantitative and qualitative methodologies, providing insight into the user community at a specific institution. Simultaneously, this research contributes to general knowledge on the patrons of small archives, a group that may be under-represented in similar studies, as the following section explains.

### **The Limits of Information on Small Archives**

There are limited data available on small museums and special library collections, and the needs of the users they serve. Indeed, a review of the literature was unable to uncover any usability study conducted at an institution comparable to the Plumas County Museum. Research in the field has focused upon larger, more prosperous institutions with the staff and resources necessary to sustain innovation (Duff, Dryden, Limkilde, Cherry, & Bogomazova, 2008). Even broad state-of-the-industry studies tend to overlook smaller institutions. Many studies rely upon association member lists to contact the archival community (examples include Walch 2006; Dooley & Luce, 2010; Dooley, Beckett, Cullingford, Sambrook, Sheppard, & Worrall, 2013; Craig, 2000; Chute, 2002;

Benedetti, 2003). Smaller institutions without the time, money, or professional-level staff to engage with professional groups are likely to be underrepresented or even absent from this type of mainstream research. In a 1982 survey of college and university archives, fewer than 50 percent of institutions reported belonging to the Society of American Archivists. “By nearly a two to one ratio, large institutions reported at least some SAA members when compared with small institutions” (Burckel & Cook, 1982, p. 416). A slightly larger number of small intuitions reported staff belonging to less expensive local or regional professional organization. Only 56 percent of small organizations reported professional staff. Large institutions reported an average 1.95 professional staff members. In the more recent A\*CENSUS study of 2006, 79.5 percent of respondents reported belonging to a professional association specific to the archival profession, though “this high affiliation...may have resulted from the process used to compile the A\*CENSUS mailing list” as “most of the names were obtained from professional association membership lists” (Walch, 2006b, p. 101).

Research populations drawn from the rosters of professional membership organizations are likely to exclude smaller organizations. Thus, while Dooley & Luce’s 2010 survey of Association of Research Libraries (ARL) member institutions indicates that 97 percent have a digital archive program in some form, the finding may not apply equally to archives and special collections libraries more broadly. Regional consortiums like the Online Archive of California (<http://www.oac.cdlib.org/>) or the Kentucky Digital Library (<http://kdl.kyvl.org/>) might serve as a more reliable indicator of digitization rates at small local archives. Unfortunately, while contributor lists are available for each consortium individually, there is no centralized cross-referenced contributor directory.

Calculating the number of institutions with digital holdings would require a comparative survey of all such sites—a major research undertaking of its own.

Archival user evaluation studies are also concentrated at larger institutions (Duff, Dryden, Limkilde, Cherry, & Bogomazova, 2008). Often, such research is designed to examine the efficacy of the arrangements at a particular institution in order to make specific improvements. Examples are Dembo and Custer's (2010) study of unique page views of the online finding aids at East Carolina University's Joyner Library and Lewis and Griffin's (2011) study of the online archival photo library at the University of South Florida. Such studies are necessarily limited to the types of organizations that can afford to conduct them and are frequently authored by staff or researchers from the same institution.

Broader studies tend to focus on selected categories of user or types of research without delving into the characteristics of the institutions that conducted research. For example, in their influential survey of historians, Stieg Dalton and Charnigo (2004) asked respondents to distinguish between discoverability tools consulted at their home institutions as opposed to those utilized at outside institutions. Participants were not asked to report on where the manuscripts they accessed were housed. Thus, while the study revealed that 94 percent of respondents considered manuscripts, archives, and special collections important to their research, questions remain. Were materials accessed online or in person? Was geographic location a consideration in determining which resources to consult? Were researchers willing and able travel to consult records? Were there trends in the size, scope, or reputation of the institutions visited? In a similar study, Tibbo (2003) examined the research practices of historians of American history at

68 premier institutions, inquiring extensively into search strategies and outcomes. The study does not, however, discuss the institutions where the sought after records were eventually uncovered. A related research project in the United Kingdom (Anderson, 2004) employed a questionnaire identical to Tibbo's. The scarcity of institutional data was also evident in Duff and Johnson's (2013) study of professional genealogical researchers. The subjects interviewed likely had experience with small and local archives, but institutional characteristics were not explicitly addressed.

Thus, generally accepted research methods are likely to unintentionally limit or exclude data on small archives and their users. Specifically, methodologies that rely on professional membership organization contact lists may tend to under estimate the number and variety of archival institutions while overestimating the prevalence of online archival programming. The lack of evidence collected from the users of small institutions makes it unclear whether their needs differ substantially from those of users of larger institutions. Indeed, information discovered on small institutions is limited to such an extent that it is difficult to report confidently on the exact number of archives within this category, or the percentage of small institutions with digital archival offerings. A potentially revealing estimate of the number of small archives comes from a recent study by the Institute of Museum and Library Services reporting that there are more than 35,000 museums in America, over half of them focused on history, with 7.5 percent classified as "History Museums" and 48 percent as "Historical Societies, Historic Preservation, and Historic Houses and Sites" (Institute of Museum and Library Services, 2014a). The vast majority of these are small or local institutions, and it seems fair to assume that many will have archival records. Culling of the supporting data files reveals



that only 12.3 percent of these institutions have websites (Institute of Museum and Library Services, 2014b). The study did not indicate what percentage of websites provide only basic visitor information, such as hours and location, as contrasted to archival records. Even at institutions like the Plumas County Museum, where digital surrogates are available for some of the more popular items, virtual copies may not be mentioned on the public website or posted online. Users can learn about digital records only by calling or writing to the institution directly.

### **Barriers to Digitization at Small Archives**

There are several reasons why small institutions may have not digitized their holdings. First, local organizations by definition serve a specialized niche audience focused upon a particular geographic area or topic. In cases where records are already available in a digitized format, posting them online requires time, effort, and computer skills that exceed ordinary word processing. In institutions with a permanent staff of one or two members supported by volunteers, other more urgent tasks may take precedence or the necessary skillset may not be readily available.

Once the materials are posted to a website, there is a regular monthly cost associated with maintaining them, dependent upon the size of the site. As of October 2015, prepackaged consumer and small business plans ranged from as little as \$4.99 per month for a GoDaddy Economy package to \$124.99 for Bluehost Dedicated Hosting Premium package. (see <https://www.godaddy.com/hosting/web-hosting> and <http://www.bluehost.com/dedicated>). Hosting costs for large sites with performance needs that include file storage and file sharing are negotiated separately on an account basis. With just one collection at 16GB, the Haun Collection test site may be considered small. Hosting fees cost \$42.00 per month. Alternatively, web designers can use a file

hosting service to store files in the cloud, where websites and other applications can reference them. Benefits include the ability to use the same file in a variety of different contexts, reduced website size, consistent performance as file libraries grow, and possible cost reduction overall. However, using code to reference external cloud-based files in this manner requires a higher degree of technical expertise than uploading files to a website directly. Amazon offers two relevant file-hosting services. The S3 service “provides developers and IT teams with...object storage” on a graduated pricing scale starting at \$.03 per GB for the first TB, with savings as usage increases (Amazon EFS, 2015, para. 1). The EFS service released in April 2015, promises an “easy to use...simple interface” for content repositories and big data storage at a stable \$0.30 per GB per month (para. 1).

### **The Move Toward Digitization**

Despite the challenges of creation and maintenance, online records have become an increasingly standard and expected part of archival practice. The pressure to bring digital records online continues to mount, even for small and local institutions. “With user expectations and options for locating information growing at an astonishing pace, special collections personnel must find solutions to the problems of access to and meaningful presentation of their collections” (Lewis & Griffin, 2011, p. 21). Many factors inform this drive. When students tackle a school assignment most begin with a Google search; many will never access non-digital sources (Purcell, et al., 2012; Kolowich, 2011). Most undergraduates “are accustomed to having information immediately available online and have little experience or patience with developing a robust search strategy” (Rose-Wiles & Hofmann, 2013, p. 148). In this environment, an online presence may open collections to an entirely new usership locally in addition to

extending its reach beyond a particular geographic region.

Likewise, for teachers, a class trip to the local museum is often impractical. The expense of travel, even across town, can be prohibitive. School bus fees vary by district but commonly include some combination of a minimum fee per bus, driver salary, and mileage charge. Estimates for a half-day trip 9:30 a.m. - 1:30 p.m. using a bus garaged 5 miles from the school to a destination 10 miles away range from \$165 to \$370.<sup>1</sup> Field trip grants fall along a similar spectrum, ranging from Big Yellow School Bus grants of \$200-\$500 administered by regional arts councils to \$700 grants offered by Target stores. In addition to securing travel funding, teachers must also account for entrance fees and meals, secure chaperones, and obtain permission slips and waivers.

Research indicates that school field trips are in decline due to budgetary constraints and increased emphasis on core competencies (Wheeler, 2011). An American Association of School Administrators (AASA) survey revealed that 30 percent of school districts eliminated field trips during the 2010-2011 school year, while 43 percent were considering it for the 2011-2012 year (2012, p. 9). “Museums across the country report a steep drop in school tours. For example, the Field Museum in Chicago at one time

---

<sup>1</sup> Price range was calculated based on publicly posted school bus pricing schemes. In Scott County School District, where Georgetown, Kentucky is located, buses cost \$21.75 per hour in driver fees, and \$1.90 per mile traveled (Scott County Schools, 2014). In Highlands County, Florida, buses cost \$13.76-18.91 per hour in driver fees, and \$1.00 per mile traveled (The School Board of Highlands County, 2012). The Los Angeles Unified School District gives a flat estimate of \$370 in bus fees per trip, noting that driver’s salary and distance may affect the cost of the trip (Los Angeles Unified School District, 2014). Akron, Ohio public schools charge \$39 per hour and \$1 per mile (Akron Public Schools, 2009). In some school districts, including the Plumas Unified School District, where the Plumas County Museum is located, as well as San Francisco and New York City, school bus pricing is accessible only to staff with a username and password.

welcomed more than 300,000 students every year. Recently the number... [has dropped] below 200,000” (Greene, Kishida & Bowen, 2014, p. 79).

At the same time, direct experience with culture is important to student development. One educational model known as constructionism theorizes, “knowledge is not transmitted but constructed with the help of other people and the support of a material environment, of a culture, of a society” (Gilliland-Swetland, 1998, p. 138). A large-scale randomized control study comparing students who visited a local art museum against those who did not found that field trip attendees scored higher in critical thinking, historical empathy, tolerance, and interest in art museums. Results indicated, “the benefits of a school tour are generally much larger for students from less advantaged backgrounds” (Greene, Kishida & Bowen, 2014, p. 85).

For students with limited opportunities to attend field trips, interacting with cultural material online may help to bridge the gap. Many museums, research facilities, and corporations now offer free virtual field trips, mediated online tours incorporating elements of written text, images, and multimedia. Students can experience virtual field trips through independent eLearning, as part of a lesson led by the classroom teacher, or via video conferencing. Many programs include lesson plans and activities to support learning. A 2005 study showed that middle school students who experienced online field trips performed better on standardized reading tests (Maryland Public Television, 2005). A comparison between 8<sup>th</sup> grade students who participated in real and virtual field trips to a nature preserve found “no significant differences” in knowledge gain (Puhek, Perse, & Sorgo, 2012, p. 173). Malkmus (2008) suggests classroom use of digitized records is on the rise at the high school level as well. “Secondary teachers are routinely using online

primary sources to teach history and other subjects. Using primary sources engages student interest, stimulates critical thinking, and invigorates classroom discussion” (p. 1).

Digital surrogates can also play a role in pre- and post-trip lessons. Orion and Hofstein (1994) showed that students perform better when field trips are integrated into the core of the unit rather than the culmination of a course of study. They recommend a 10-hour preparatory unit leading up to the trip followed by a summary unit. “Pre-visit activities provide prior knowledge that can aid in the understanding of experiences at the site, while post-visit activities strengthen new connections and give additional context for future experiences” (Anderson, Kisiel & Storksdieck, 2006, p. 366). Online archival records may provide teachers with useful tools to incorporate into pre- and post-trip instruction.

In high-level research, new methodologies and changes in historians’ search practices make smaller, lesser-known collections potentially more relevant. Since the 1970s, historical scholarship has increasingly trended toward social history, which focuses upon the lives of ordinary people, analyzing historical events within the broader context of prevailing culture, and conceiving change as the result of zeitgeist rather than the machinations of a few powerful people at the highest levels (Herubel & Goedeken, 1993). Within this view of history, primary source material that might formerly have been seen as insignificant can take on new importance. In some cases, holdings at small and local institutions may have relevance beyond the boundaries of the communities they were founded to serve. In light of historiographical shifts, all records should be periodically reassessed.

In addition, growing interest in big data and increased access to data mining technologies could allow for comparison of aggregated digital data in the future. As Lockers, Sag and Schultz (2012) point out,

Advances in computer technology combined with the availability of digital archives are allowing humanities scholars to do what biologists, physicists and economists have been doing for decades—analyze massive amounts of data. (p. 29)

The humanities research community is just beginning to explore potential data mining applications through projects like Big UK Domain Data for the Arts and Humanities (<http://buddah.projects.history.ac.uk/>), a collaboration of the British Library, the Institute of Historical Research, University of London, the Oxford Internet Institute, and Aarhus University. In the hard sciences, there exists a more established vision in which researchers openly share their own datasets and in turn are able to access, validate, cross-reference, and process datasets produced by other groups (Pampel & Dallmeier-Tiessen, 2014). It is possible to imagine a future in which digital archives become a source of computational historical data and the many incidental records stored at small and local institutions combine to form a rich, nuanced view of the past.

A recent rise in the popularity of ancestral family research and a corresponding increase in online tools to assist this process may also have positive consequences for local collections. Records of only minor significance to historical study generally may be highly relevant to family researchers anxious to place a particular person in a particular location in a particular period. Given the rise in mobility over the 19<sup>th</sup> and 20<sup>th</sup> centuries, many ancestral researchers no longer live in close proximity to their progenitors, making online records desirable even for skilled researchers. An online presence is essential in order to reach inexperienced researchers who rely upon online tools like ancestry.com.

Barth's (1997) research emphasizes the importance of automated records for genealogical researchers, recommending online discovery tools as a way of promoting institutions to a broad user base, thus ensuring long-term survival. Duff and Johnson's more recent (2013) study of genealogical researchers bears out earlier predictions. All participants reported consulting web-based finding aids, though the specific types of finding aids were not discussed. In the context of this evolving landscape, archival holdings at small local institutions may conceivably have a wider reach than traditionally imagined. Positive outcomes are not restricted to increased access. Genealogical communities can offer strong support to the archives and special libraries that serve them well, through political advocacy, donations, and volunteer labor (McKay, 2012).

As previously discussed, digitization, like other forms of facsimile that preceded it, functions as a practical solution to issues of preservation and access. The Association of Research Libraries recognizes digitization as an acceptable preservation strategy (Conway, 2010). Digital surrogates extend access to users beyond the physical reading room without risk to the original. Even within the archive, a digital copy protects the original record from damage and deprecation through handling and provides a small measure of insurance in the event of a larger disaster (Arthur, Byrne, Long, Monitori, & Nadler, 2004; Hughes, 2003, p. 11). Online offerings are also emerging as a form of marketing and community engagement, an element in maintaining the vitality and relevance of organizations, and by extension, a factor in securing funding. Despite the dangers of equating usage rates with long-term value for individual records, access and use are nevertheless a major consideration for funders. Donors and resource allocators demand accessibility (Meissner & Green, 2010). If web-based offerings increase access,

they have the potential to increase funding as well. Further, while digitization efforts place heavy demands on budgets, they also justify grants, as illustrated by the myriad of grant offerings available in the United States from organizations like the National Endowment for the Humanities, Council on Library and Information Resources, and the Institute of Museum and Library Services at the federal level, as well as numerous state and regional agencies and charities.

### **The Research**

This research explores questions of online archival presentation for small and local collections. During 2013 and 2014, digital surrogates were created for a collection of diaries, letters, and ephemera housed at the Plumas County Museum in Quincy, California ([plumasmuseum.org](http://plumasmuseum.org)). The resulting images were posted to a test site, [hauncollection.org](http://hauncollection.org), in three different presentation versions. The first version represented the industry standard with scanned images arranged to approximate the experience of visiting the physical archive. A second intermediate version paired scanned images with transcribed text and some limited cross-referencing between records. Finally, a third more interpretive, curatorial version used the flexibility of the online environment to create a networked, highly integrated experience enhanced by additional resources including annotation, biography, a map, and other interpretative tools not typically available. A variety of users, including community members, museum stakeholders, scholarly researchers, and genealogists were invited to explore the test site and provide feedback about the usefulness of the information provided in a short survey. Data concerning user behavior within the site were gathered using Google Analytics. Results provide insight into the questions surrounding the digitization and presentation of archival content at small institutions: Does an online presence extend the user population



for local collections? If so, do digital collections simply attract greater populations of the same types of users who frequent the physical reading room, or do they open the collection to new groups accessing the material for the first time? How do users prefer to interact with primary source material within a digital environment? What features do users find most useful? Do factors such as research goals, past experience in archival research, or level of education have any impact on preferences?

### **Thesis Organization**

The study is organized into five chapters. Following this introductory section, Chapter 2 provides a survey of related literature, paying particular attention to emerging digital standards and the challenges they carry with them. The third chapter explains the study design and methodology, including background information on the institution and collection used in this test, an account of the site construction process, tester recruitment, and data collection. Chapter 4 presents the results of the study, analyzing collected data and discussing its implications. The final chapter places the outcomes in the broader context of the field and discusses the need for future study.

## Chapter 2: Literature Review

From the first appearance of digitized finding aids in the early 1990s, online archival resources have occupied a prominent but highly contested position within the profession. The literature review explores both the emerging norms of digital archival practice and the practical and ideological challenges presented by the online imperative.

### Emerging Norms

Online archives are currently offered in a broad range of formats. Some adhere to established archival standards, seeking to approximate the experience of visiting the real archival institution as closely as possible. In this approach, records are arranged according to series, box, and folder, with a descriptive finding aid provided to help guide the user. Often, these sites develop over time as institutions systematically digitize existing finding aids, mapping them to scanned images as they become available. This approach is often the easiest from the organization's perspective and, some argue, most appealing to the traditional archival user, the experienced scholarly researcher.

At the opposite end of the spectrum are the sites whose creators have taken an interpretive route, drawing inspiration from other disciplines to deliver a more structured, contextualized experience to the end user. As Lewis and Griffen (2011) have observed, "Many libraries with digital collections are adopting the concept of themed or contextualized Web sites that provide background information, related resources, scholarly essays, and similar types of material to inform and orient their visitors" (p. 22-23). This approach, which emphasizes the curatorial and interpretive aspects of the archival profession, is also known as "creative archiving" (Velios, 2011) and "interactive archiving" (Daines & Mimer, 2009). For example, the "Documenting the American South" (<http://docsouth.unc.edu>) project at the University of North Carolina at Chapel

Hill, presents “selected documents available alongside contextualizing documentation from archivists and scholars” (Louis Round Wilson Library, 2009). Records are organized into thematic collections, as well as by author, subject, title, and location. Additional resources include K-12 lesson plans, maps, and written narrative. The John Latham Archive ([www.ligatus.org.uk/jla/node/74](http://www.ligatus.org.uk/jla/node/74)) at the University of London is particularly noteworthy, as it offers standard fields for archival description, a randomized record viewer for browsing, and an alternative organizational system drawn from the artist’s own theories and philosophy.

Online archives in general, and the creative approach to archival digitization in particular, are contentious both because they are much more demanding than straightforward processing and because they have not always found favor with experienced archival researchers. In a series of needs assessment interviews conducted in 2009, researchers were surprised to find their subjects expressed wariness about the reliability and credibility of web 2.0 features and frustration with themed collections (Allison-Burnell, Yakel, & Hauck, 2011). Proponents of online archiving argue that both the business of archives (i.e., the sources of and justifications for programmatic funding) and the potential user population are changing and that the profession must change in response to remain relevant.

When expectations for digital access reach the point where only digital information will satisfy the vast majority of user needs—as may already be the case in universities, corporations, and government agencies—then libraries, archives and museums must digitize extensively and to a level of quality that supports a wide variety of actual and potential uses. (Conway, 2010, p. 74)

Many organizations straddle the line between established norms and full commitment to creative archiving. In some cases, this is the result of an organic process.

Early adopters may have begun by mimicking physical processes in an online environment, moving on to experiment with new forms as time wore on, while still maintaining existing resources. In other cases, taking different approaches has been a very deliberate decision designed to appeal to specific user groups with different needs. For example, the Online Archive of California (<http://www.oac.cdlib.org/>) is aimed at university students and serious researchers, while its sister site, Calisphere ([www.calisphere.universityofcalifornia.edu/](http://www.calisphere.universityofcalifornia.edu/)), is geared toward K-12 students and the general public. Other popular examples of the hybrid online archives take varied approaches but, importantly, all offer multiple ways of accessing, viewing and understanding content. The Franklin D. Roosevelt Presidential Library and Museum ([www.fdrlibrary.marist.edu/archives/collections.html](http://www.fdrlibrary.marist.edu/archives/collections.html)) offers material through both a database of digitized finding aids, such as lists of manuscript collections and archival folder titles, and a more flexible keyword search of digitized records available online. The Library of Congress' extensive "American Memory" collection ([memory.loc.gov/ammem/index.html](http://memory.loc.gov/ammem/index.html)) consists of 140 digitized collections that may be browsed by collection title or thematically by topic, as well as searched at the item level via a simple in-browser search field. The "Polar Bear Expedition Archive" ([quod.lib.umich.edu/p/polaread/](http://quod.lib.umich.edu/p/polaread/)), heralded as an exemplar by researchers in the area, similarly offers several different approaches to its digital collection: a traditional hierarchical finding aid starting at the collection level; an index style search option for person, place, Army unit, or geographic location; a complete item level listing; and a narrative history.

## **Challenges of Theory and Practice in the Field**

While the digitization of physical records is prevalent and digital originals will soon outpace physical records entirely, several significant areas of tension within the archival field still surround the issue. They include 1) principles of archival arrangement and representation; 2) archival processing standards and practices; 3) the question of usership and; 4) issues of research quality and access.

**Archival arrangement and representation.** Digital archives, with their selectivity, inherent flexibility, and heightened capacity for cross-reference “...represent an expansion of the boundaries of traditional archives” (Monks-Leeson, 2011, p. 39) and challenge the fundamental concepts of provenance, original order, and objective presentation. The archival principle of provenance dictates that material created or received by a single individual constitutes a collection, and further, that collections should not be artificially divided or combined by the archivist (Zhang, 2012, p. 49). The severing of one segment of a collection or the integration of material of varied provenance into a selective composite collection is negatively viewed by the profession as editorial license, drawing new and potentially biased intellectual links, and placing the material in a novel context that might potentially alter its meaning and significance. Synthetic composite and sub-collections are thus generally considered permissible for temporary curated exhibitions, publications, or online offerings, but not as a permanent adjustment to the collection itself (Westbrook, 2002, p. 75). Digital objects are routinely presented selectively in the online environment, outside the context of their collections, in the company of material from different collections and even different institutions. These exhibitions may provide a limited amount of metadata expressing each item’s original provenance or the scope of the larger collection. This may not have greatly concerned

researchers when digitization was the exception and all digital collections could be considered in some sense “special exhibitions.” However, as digital records become the standard—and for many users the sole—means of access, the implications are far more sweeping than they may have first appeared.

The widely accepted principle of original order is also compromised by current trends in digital archival presentation. This principle stipulates that the arrangement of archival records maintain or replicate as nearly as possible the arrangement of the records while they were in use (Hunter, 2004, p. 114-115). Original order relies on the assumption that the arrangement of records when active offers “insight into the record creating processes” and the ways in which the records were referenced and used. The profession views original order as an important strategy for assuring authenticity and objectivity in archival arrangement, a way of organically preserving and documenting the records in their natural state (Zhang, 2012, p. 52). Historians value original order as contributing to a deep and contextualized view of the subject matter, conveying significant implications about the relationships between the records and the constructions placed on them by their creators, users, and preservers. To borrow an example from Breakell (2010) at the University of Brighton Design Archives,

with minimal information to show the structure of the files, the gradual unpicking of the inherent structures of the archive was almost a forensic process...the overall structure and shape of the archive reveals a further layer of information about how the institution saw its functions, its concerns and priorities; and indeed what it felt should be preserved for its own purposes and those of the notional future historical. The total evidential value of the archive is more than the sum of its component parts. (p. 13)

Original order combined with a holistic view of the collection also highlights gaps in the records, which may have significant implications to research. One of the historians

interviewed in Duff and Johnson's 2012 study articulates this, asking, "when a gap of four years suddenly appears in a person's diary collection, is it because 'the person is grieving for four years, or because there is a fire?'" (p. 487).

In digital archives, however, original order is frequently lost or bypassed. Partial or selective digitization obliterates the delicate relationships between items. Even when the entire collection is present, the search, tagging and cross reference capabilities which are among the best arguments in favor of digitization also result in the creation of multiple points of entry, generating "new contexts for textual, image and media records" (Monks-Leeson, 2011, p. 39) and inviting new perspectives on the material (Breakell, 2010). Many of these changes are unwitting and unavoidable functions of how digital environments work, but there has also been a more deliberate shift in the way archival institutions present information independent of both provenance and original order. Digital archival material is now frequently arranged on a flexible thematic basis by historical period, subject or record type rather than original order (Monks-Leeson, 2011). The increased power of and reliance on search engines means that in many cases users will access records directly from Google, completely bypassing the contextual framework of the larger collection (McCausland, 2011). This narrower focus, some argue, contributes to a more shallow understanding of the material and may have serious consequences for historical practice (Newell, 2012, p. 289).

### **Processing standards and practices**

In addition to the tangled issues of theory and principle, there is also the practical question of how to get archival material online. Despite a great deal of interest and effort, archives have lagged behind the rest of the library world in the move toward the

digital. In part, this is due to the added complexity of having to digitize paper records as opposed to acquiring born-digital resources—but there is much more behind this discrepancy than the challenges posed by scanning and photography.

In the influential 2005 article that sparked the More Progress, Less Process (MPLP) movement, researchers Green and Meissner highlight the serious processing backlogs plaguing archives. In addition to their own research, the authors cite a 1998 survey by the ARL indicating that unprocessed material at the time amounted to nearly one third of all collections. In light of their findings, Greene and Meissner argue that archivists should shift their focus from item-level to collection-level processing, dispensing with detailed arrangement as well as preservation tasks, such as removing staples and refolding, in order to maximize the amount of material accessible to the public.

The MPLP model has been received with enthusiasm in the profession. In a 2010 survey by the OCLC, 75 percent of respondents indicated that they use an MPLP approach, 57 percent sometimes, and 18 percent always (Dooley & Luce, 2010). In addition to a very practical shift in processing standards, MPLP hints at an implicit underlying shift in values, a re-envisioning of the archival profession in which the needs of the end user are considered on par with the needs of the collection (McFarland, 2007). This shift is apparent in staff allocation as well. As the OCLC researchers warn, saving time in processing does not mean saving time overall; less processed collections may impede discoverability, requiring staff to devote more time to engaging with and supporting researchers (Dooley & Luce, 2010, p. 49). McFarland (2007) makes the same



point with a slightly more positive spin, positioning increased public engagement as one of the great benefits—even one of the goals—of an MPLP system.

Green and Meissner’s recommendations have lately been called into question by Prom (2012) whose reanalysis of their data indicates that the level of processing may not be the real cause of the serious backlog in archival processing. His findings suggest, “there is no statistically significant correlation between use of intensive processing techniques and slower processing speed” (Prom, 2012, p. 157-158).

The addition of digitization to the standard processing workflow undeniably increases already unsustainably slow processing times. The drive toward mass digitization may be viewed as both an extension and a contradiction of Green and Meissner’s widely accepted logic. Digitized records offer greater potential opportunities for end user access but also require archivists to engage more fully with records, handling each item at least long enough to create a digital surrogate, label it meaningfully, and upload it to the appropriate location. Similarly, “tension exists between many digitization initiatives that create digital collections with robust item-level metadata and context, and the traditional archival manner of defining and describing collections at the aggregate or group level” (Gueguen, 2010, p. 96).

It remains unclear where the greatest benefit in terms of accessibility lies. Digitization might conceivably facilitate access to archival collections by large numbers of new users. However, the demands of digitization will likely mean increased backlogs, resulting in a smaller percentage of collections being accessible through any means, digital or analog. A clearer understanding of use patterns and the perceived benefits of

digital archives by the user population is necessary in order to inform this cost-benefit analysis.

**The question of usership.** The demand for more information about user behavior is not new. In an influential 1984 issue of *The American Archivist*, Elsie Freeman questioned whether archivists adequately understand user identity, the nature of patrons' research, and the types of support required. In the same issue, William Joyce (1984) made a similar case for leveraging the expertise of the scholarly community to improve archival practice, while an editorial from genealogist Mary Speakman (1984) offered a sample of the type of feedback archivists might expect from researchers. The issue as a whole and Freeman's call for needs-driven programing and systematic nationwide user surveys to more clearly define the user population and its needs struck a chord with other researchers. Jacqueline Goggin cited Freeman's work in her own 1986 study of scholarly use patterns. Tracking the published works of researchers who accessed 13 collections on women's history and black history, she found that the primary source material had been used sparingly in many cases. She speculated on the possible reasons, citing a lack of processing, limited descriptive resources and finding aids, and unarranged files as possible factors. Goggin (1986) argued that reference interview records and access registries be directly employed in processing strategy in order to ensure that the material in high demand is available. Dowler (1988) took this a step further, advocating for records evaluation, and by extension program development, that considers potential as well as actual use. He encouraged archivists to consider the needs of "future users and all those who could use, might use, perhaps even should use, the information in archives" (Dowler, 1988, p. 78).

This type of thinking spawned numerous studies and surveys in the decades that followed. Researchers have examined scholarly search behaviors, including structured interviews with users of specific collections or institutions (Duff & Johnson, 2002, 2003; Yakel, 2002); formal needs assessments (Allison-Burnell, Yakel, & Hauck, 2011); user surveys administered both onsite and online (Krause, 2007); studies of archive professionals' perceptions about user studies (Duff, Dryden, Limkilde, Cherry, & Bogomazova, 2008); quantitative analyses of researcher search queries and reference requests (Bantin & Agne, 2010; Yakel & Bost, 1994), repository site traffic (Dembo & Custer, 2010), and citations from the published scholarly works of researchers who visited the archive (Groggin, 1886); broad cross sectional studies in which users were asked to describe their own search strategies and research habits (Stieg Dalton & Charnigo, 2004); and studies that examine and evaluate the various methodologies and research models available (Snow et. al., 2008). Not infrequently, two or more methods were combined in a single study. The result is a complex and piecemeal landscape, a patchwork map drawn to many different scales, defying meta-analysis and even in many cases meaningful comparison. The gap Freeman (1984) intuited between archivists' assumptions about their users and the reality has become increasingly evident.

Archival practice, from appraisal to processing to reference, rests on the assumption that archives primarily serve a captive audience of scholars equipped with strong research skills. The research, however, indicates that scholars are not the only audience for archival collections—sometimes they do not even comprise the primary audience. For example, Breakell (2010) reported that much of the interest in archival digitization is driven by family genealogical research. According to Bantin and Agne

(2010), students also make heavy demands on archives, with “large numbers of first-time researchers looking for quick answers to questions on popular topics” (p. 245). The population of lay historians seems to be on rise, perhaps linked to the baby boom generation leaving the workforce (Hatton, 2006; Showley, 2007). Nor are scholars required to visit archives. Conventional wisdom dictates that because their content is unique, archives have an effective monopoly, a built-in audience of related researchers who cannot go elsewhere. Studies of scholarly research methods refute this assumption. Goggin (1986) suggests that the historians’ lack of training in research methods and the intimidatingly large volume of available resources may discourage scholars from using archival sources in their research, indicating that the “experienced scholarly researchers” may not, in fact, be as experienced as had been previously assumed (p. 62). In 2002 Yakel made a similar observation, “Archivists may also be assuming that users understand more about archival operations and archival terminology than is warranted. As a result, archivists may be overestimating the expertise of users and their ability to transfer knowledge from one repository to another” (p. 112).

In common with other types of archival research, the majority of user studies discussed here were conducted at large archives and museums or within the context of major research universities—often the author’s home institution. For example, Yakel’s (2002) work was conducted at the University of Michigan, an R1<sup>2</sup> research institution, while Allison-Bunnell’s (2009) study addressed the development of the Northwest

---

<sup>2</sup> In the Carnegie Classification of Institutions of Higher Education framework (<http://carnegieclassifications.iu.edu/>) R1 institutions are those classified as having “very high research activity.” Other classifications for research institutions include R2 “high research activity” and R3 “doctoral/research universities.”

Digital Archives, the product of a long-term collaboration between more than 40 academic institutions across five states. As Duff and colleagues have observed, “user-based evaluation studies tend to be conducted in larger institutions, and often the findings are not widely distributed” (Duff, Dryden, Limkilde, Cherry, & Bogomazova, 2008, p 146). For this reason, while the peer-reviewed literature establishes methodological norms and important insight into the issues users face, outcomes are not directly comparable to this study of a small, local institution.

**Impact of digitization on research quality.** Digitization has been characterized both as a solution to enhance access to archival records and a threat to rigorous research (McCausland, 2011). Digitation proponents assert that online archives allow for “new combinations of text, images, and ideas to inspire research as well as the ability to conduct different forms of data mining and analysis” (Zastrow, 2013, p 16). An online presence equates to unlimited ubiquitous accessibility from almost anywhere in the world using a range of increasingly affordable devices. Not only does online availability enable traditional archival users in remote locations to access materials, it has the potential to open collections to new user groups who might not otherwise have become aware of their existence. For example, in December 2014 the San Diego Air and Space Museum reported an increase in online commerce after it began uploading 135,000 digitized images from its collections to Flickr. In fact, the museum’s Flickr page received over four million views in its first 18 months (San Diego Air and Space Museum, n.d.).

On the other hand, archivists, like library professionals from all specialties, are troubled by the popular assumption that everything important is now publicly available—or at least mentioned—on the Internet. The common misconception that all the world’s

knowledge is just a Google search away is seen as undermining thorough research practices, leading some researchers to ignore older resources, smaller journals, and vast swaths of primary source material. Today “we’re starting to see researchers who don’t have the time, travel funds, or inclination to seek out materials they can’t get online” (Zastrow, 2013, p 17). Thus, “while our collections are unique, we should not delude ourselves that the majority of our potential users must come to the archives” (McFarland, 2007, p. 142). Even among users who are well aware that not all collections are available digitally, there is an increasing expectation that they will be, or should be.

McCausland (2011) nicely summarized archivists’ concerns over unmediated or unguided access to collections. She reported that unlike libraries, archives are not designed to serve independent researchers. A finding aid “places the materials in context by consolidating information about the collection, such as acquisition and processing; provenance...scope of the collection...organization and arrangement; and an inventory of the series and the folders” (Society of American Archivists, 2015). Finding aids are not equivalent to the catalog entries used in self-service libraries and rarely offer item-level listings or descriptions (Hunter, 2003). Similarly, there is no serendipitous browsing permitted; patrons do not stroll through the stacks, poking into the boxes that happen to catch their eye. Instead, visitors to an archive interact with the material through the mediation of the archivist. Archivists facilitate research through discussions with patrons. They recommend resources and provide supervised access to material one box at a time.

Mediated access has evolved over time out of practical necessity. Many archival materials are delicate enough to preclude unsupervised access. Collections tend to be

large, complex, and interconnected, making it difficult for anyone unfamiliar with the records to navigate them efficiently, whether they are a skilled researcher or not (McCausland, 2011). Collections that have received MPLP compliant minimal processing may be especially opaque, even to the archivists who processed them (Van Ness, 2010). In the case of students, and presumably other inexperienced researchers, there is a risk that individual records may be viewed out of context and misinterpreted (Gilliland-Swetland, 1998). Online archives require not only digitized surrogates but also an entirely new system for organizing and categorizing records. Simply posting a preexisting finding aid and associated scans by series may not be sufficient for every collection, and with no reference archivist available for consultation, records may remain as inaccessible as if they had never been posted online at all. “Archivists must find a way to capture systematically the knowledge of the reference archivist and enter this information into finding aids and knowledge-based systems for providing access to records” (Dowler, 1988, p. 82). Further, “users need to do more than locate information in archives; they also need assistance to ask questions and make sense of their results” (Johnson, as quoted in McCausland, 2011, p. 315).

### **Conclusions from the Literature**

Since digitization and online archival presentation first emerged as methods of preserving delicate records, extending access, and handling the new wave of “born digital” records, many varied applications have been implemented. In early days of archival digitization, Gilliland-Swetland (1998) observed that online archives had so far developed as “individual digital access initiatives that are rarely fully articulated, systemized across repositories, nor designed based on an analysis of users and their

needs” (p. 142). More recently, Conway (2010) argued for consistent, blanket digitization for preservation, asserting that we are at the end of the era of “boutique digital scanning projects for which the principal goal is experimentation with new technologies” (p. 76). Models range from an archival approach, which seeks to uphold established archival values and practices, replicating the physical archive in an online environment, to a more malleable, thematic, and often-personal approach, which embraces the variability and flexibility of the online space and forces archivists to reinterpret foundational principles of the profession and their own roles.

A rich body of peer-reviewed research explores both the practical aspects of online archival development and the intrinsic theoretical tensions inherent in the shift from analog to digital archival interfaces. The literature explores the continuing importance of archival arrangement in understanding records and considers whether the sacrifice in contextualization and the “provenancial bonds that archivists take as crucial to a record’s meaning and evidential value” is worth the benefits of online access (Monks-Leeson, 2011, p. 39). Similarly, many writers in the field express concerns about the quality of the research performed online. Are quick answers necessarily less thoughtful ones? Will the less accessible records simply be ignored? What does selectivity based on convenience mean for scholarship and the study of history? These issues are more immediately relevant to the traditional archival audience of scholarly researchers than to the more recent and numerous users among the public, but clearly, the answers will shape our collective understanding of the past. It seems probable that cross-disciplinary collaborations with researchers in historiography or even psychology could offer speculative answers to some of these questions.



In the area of practical application, ubiquitous online access demands processing standards (including standards for handling records, creating digital surrogates, and generating metadata for discoverability) far beyond what even the most preeminent archival organization can possibly perform and in direct contrast to prevailing MPLP wisdom. The situation seems to cry out for a technological solution that will further automate the process. Until then, archivists are forced to be selective.

Finally, various studies have sought to address user needs. Some examine the success of particular programs within specific organizations. Others investigate researcher habits and the outcomes of research. The needs of historians appear relatively well understood through the work of Stieg Dalton, Tibbo, Anderson, and others. Further investigation into the institutions where historians conduct this research would, however, offer useful insight for archivists trying to assess the relevance of such studies for their own organizations and programs. In addition, further study of historians' research strategies when working with physical records versus digital facsimiles would be valuable. For example, do researchers spend the same amount of time reviewing a record regardless of format? Do notetaking and conclusions drawn differ according to the format of the record consulted? Are researchers more likely to cite records accessed in person rather than online? The results of an A/B<sup>3</sup> study of a group of researchers of similar expertise and experience would be fascinating.

Far less has been reported on the needs and information-seeking habits among archives' other users, though as Anderson (2004) points out, there is a "growing

---

<sup>3</sup> A/B testing is randomized study comparing two alternatives common in user experience design validation.

recognition of the diversity of users” (p. 85). Beyond the scholarly community, who consults archival records, what do they hope to learn, what are their expectations upon entering an archive or accessing an online record, what drew them to the material in the first place, and how will they measure the success of their visit? Does the relative composition of user population vary according to format, collection, subject specialty, institutional size and reputation, proximity to urban centers, or other factors? What trends emerge? Do these patterns differ in an online environment? The scope for investigation here is boundless.

This study examines just a few of the myriad questions surrounding digitization through the lens of one collection, housed at the rural Plumas County Museum. Specifically, the work aims to gauge the value of online records for small and local institutions, exploring the scale and composition of the online audience for this niche material, and simultaneously determining which of the popular approaches to web-based archival presentation best serve the needs of these users. The outcomes will have meaningful implications not only for the Plumas County Museum and the more than 200 other local historical museums and associations in California, which may well share a highly similar user base, but for small archival programs more broadly. Given the limited information available about online archives at small and local institutions, this study presents rare insight into an underrepresented population. Similarly, little has been written about archival users outside the formal research or academic library setting. By including a broad spectrum of small archival users, the present study explores patterns in user preference and need across multiple user types, as well as the nuanced relationships between user groups.

## **Chapter 3: Study Design and Methodology**

### **Research Setting**

The Plumas County Museum was founded in 1971 for the purpose of “preserving, protecting and promoting the history of Plumas County” (Plumas County Museum, n.d.). The institution serves a small, rural community of 20,007 spanning 2,613 square miles of mountainous terrain in the northern Sierra Nevada range 60 miles north of Lake Tahoe (Plumas County, California, n.d.). Public areas include a central exhibition room and mezzanine with displays of Maidu basketry, taxidermy, historic toys, weapons, medical equipment, skiing paraphernalia, biographies of local historical figures, Chinese artifacts, and more. Outbuildings feature geological and timber exhibits, a working smithy and gold panning area manned by docents, a carriage house with vehicles (including a 1922 Studebaker currently undergoing restoration by the auto shop class from the local high school), and a miner’s cabin relocated from its original site. Two adjoining properties, the Variel home (a restored 1878 house open for tours) and the Goodwin Law Office (constructed in 1859) also belong to the museum.

Extensive storage space houses donated items including preserved papers, ledgers, county documents, photographs, clothing, and artifacts. A reading room offers reference works and files on county residents, places, and organizations. Files contain not only donated items but also research conducted by museum staff, family trees, transcribed interviews with area residents, news clippings, and copies of related records at other institutions.

The Plumas County Museum attracts over 5,000 visitors per year, with an average of 5 guests per month or 60 per year who conduct in-depth research in the reading room. The museum employs a full-time director and part-time assistant supported by an active

volunteer group. Five directors appointed by the county Board of Supervisors and an 11-member Board of Trustees, including the director, make up the organization's governing body. The director's funding comes from Plumas County. All other funding, such as operating costs and the assistant's salary, comes from the Plumas County Museum Association, a membership organization.

Though the museum does not routinely post digitized records online, many items, particularly photographs, have been previously digitized and are available to users upon request. The museum offers photocopying and scanning services for a small fee. The organization's longstanding practice when reproducing items for patrons has been to make a preservation copy for its own records.

### **The Haun Collection and Reasons for its Selection for This Study**

The archival records selected for the test, referred to as the Haun Collection, is comprised of diaries, letters, and ephemera covering more than 100 years in the life of one family. Four of the five Haun brothers and their families migrated from Kentucky to the California gold fields between 1849 and 1855. Attorney Henry Peter "H.P." Haun and his wife and cousin Catherine traveled overland during the original 1849 rush. The couple settled on a ranch outside Marysville, California, where H.P. practiced law before being elected as a judge. A younger brother, Andrew Jackson "Jack" Haun joined him there and worked for him on the ranch. James Haun and his teenage son John followed by sea in 1853 and settled near Nelson Point on the Middle Fork of the Feather River, where they amassed a small fortune. The youngest Haun brother Dave accompanied James and John aboard ship, spending time in both Marysville and the gold camps before returning to Kentucky to finish law school. James' wife Martha Haun remained behind

in Georgetown, Kentucky, winding up the family milling business, caring for an orphaned niece, and overseeing the family's four adult slaves and their children. Martha and her niece joined the men in California in 1855, traveling under Dave's escort. She left the slaves behind in Kentucky, where they were hired out as servants. Shortly after his wife's arrival, James Haun purchased the American Ranch in Quincy, which the family operated as a farm and hotel.

In 1861, John Haun returned to the family home in Kentucky, where he joined the Confederate cavalry and became one of General Morgan's raiders. He was captured with the General in Ohio and spent 18 months as a prisoner of war at Camp Chase. Upon his release, he married longtime sweetheart Mollie Burns of Georgetown, Kentucky. The couple later returned to California and settled in Plumas County.

Both James and John Haun wrote daily journal entries during the gold rush and preserved much of their correspondence. In later years, granddaughter Birdina ("Birdie") Haun carried on that tradition. Birdie was also responsible for preserving the family papers and donating them to Plumas County Museum along with her own personal diaries and letters, the latter of which remain to be processed.

The Haun Collection has been organized by format and chronology into boxes and folders accompanied by hand written finding aids or notes. For the purposes of digitization, the collection was further divided into series and a formal finding aid created. Series I consists of six volumes of journals written by James and John Haun between 1853 and 1859. They discuss the journey west, life in the gold camps, politics of the day, brushes with Indians, and the gradual transformation from the gold rush economy to an agrarian one. Series II includes correspondence from 1841-1860. These

letters are primarily from Martha to her husband and son and discuss her financial arrangements, Georgetown news, and her anxiety to be with her family again. Series III covers the Civil War and its aftermath from 1861-1867. Letters between John Haun and his sweetheart Mollie Burns dominate this series. Readers watch as their relationship evolves, gleaning details of life in Civil War Kentucky and Camp Chase prison. Series IV covers 1869-1900 and finds John and Mollie married and residing in Quincy with their five children. The correspondence from this period includes letters from absent friends and family as well as letters from Mollie to John during an extended trip home to Kentucky. Finally, Series V contains cards and letters sent between 1900-1931, primarily to an aging Mollie and to her youngest daughter Birdie.

The appeal of such a collection for local history enthusiasts is clear. The Hauns were a prominent area family. James Haun owned an important ranch and lodging house. His younger brother David was a lawyer and operated a sawmill in nearby Greenville. He also served in the state legislature in 1860 and held the office of Plumas County District Attorney from 1870-1973. Local landmarks still bear the Haun name and the family were personally acquainted with the progenitors of many families who still live in the area. The Haun's story is an iconic one for the region, representative of the experiences of many early settlers.

The collection's potential research application is also evident. The extended time period covered, as well as the specificity, breadth, and universality of the experiences described, makes this an ideal exemplary case—a potentially useful resource for high-level research by scholars in economic history, California history, history of the West, women's history, and Kentucky during the Civil War. Select passages may interest

specialists focusing on Native American history, African American history, and the history of the Chinese in the gold rush. The documents can also serve as a first introduction to primary source material for undergraduates or high school students.

Finally, because the Haun family often recorded gossip, describing weddings, deaths, fights, news from absent friends, and other local gossip, the collection can provide invaluable missing pieces for genealogical researchers interested in comings and goings in the gold camps or which Confederate soldiers took the oath.<sup>4</sup> This is especially true given the general difficulty in tracking particular individuals through the turmoil of 1850s California and the Civil War and the loss or destruction of many local records. The Scott County Courthouse, along with all its records, was destroyed by fire in 1837. No Plumas county newspaper survives before 1856. In some cases, these records may not be merely a useful resource but the only resource.

In light of these considerations, the Haun Collection has value for many researchers and makes an ideal test case for this study. In formulating the study's methodology, three main user groups were identified:

- 1) Residents of Plumas County, comprised of existing museum patrons and new users with similar background who are interested in local history and in the sense of community that springs from it. Many will have firsthand knowledge of the people and places discussed in the Haun papers.
- 2) Local Plumas County educators and their students, including both high school and junior college populations, especially students of state history or American History.

---

<sup>4</sup> During the Civil War prisoners were required to take an oath of loyalty to the United States of America before their release or as a condition of presidential pardon.

This group has some overlap with the first, but their unique needs deserve separate consideration and analysis.

- 3) Academic researchers interested in topics such as the history of the West, California history or the gold rush; American migration; 19<sup>th</sup> century economic history; or Civil War-era Kentucky.

Additional potential audiences may include local history groups in Scott County, Kentucky, where the Haun family lived before migrating to California and which the women of the family in particular, discuss at some length in their correspondence.

Finally, the collection may interest relatives of the Haun family and their associates, as well as genealogical researchers more generally.

### **Test Environment**

As a test site, hauncollection.org was not conceived or intended as a visual representation of a production-ready digital archive site. Instead, the site serves as a functional model offering sample options for evaluation and should not be viewed as a final and complete resource. Nevertheless, developing the test environment was crucial to the study.

Version 1 represents the current archival standard for presenting digitized collections. It includes high quality digital reproductions presented in original order and navigable using a series-level finding aid. Users click the red links to open a high-resolution image of each page separately in a new browser tab, as shown in Figure 1, images available for download.





## Series II, 1841 – 1860

try [Version 2](#) | try [Version 3](#)

- June 30 1841 – Elizabeth Hurst to Martha Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- March 21 1843 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#)
- February 1 1853 – Annie to unknown | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- February 2 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- February 24 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- March 15 1853 – WH Haun to John J. Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- April 2 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- May 24 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- June 8 1853 – Martha Haun to John J. Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#)
- June 27 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#) | [Page 7](#) | [Page 8](#)
- July 18 1853 – Martha Haun to John J. Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- November 29 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- December 20 1853 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- January 6 1854 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- January 15 1854 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- January 22 1854 – Martha Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- February 1854 – Lizzie Hurst Haun to James Haun | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)
- February 19 1854 – promissory note | [Page 1](#) | [Page 2](#) | [Page 3](#) | [Page 4](#) | [Page 5](#) | [Page 6](#)

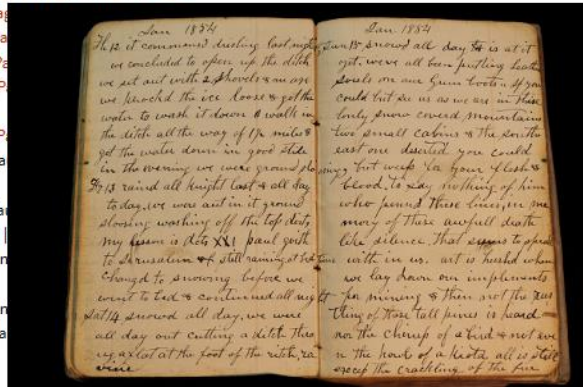


Figure 1. An example of the version 1 presentation with inset view of an individual record. Reprinted with permission of the Plumas County Museum Association.

Version 2 uses a hybrid format. It provides high quality digital reproductions presented in chronological order paired with a verbatim transcription, preserving errors and misspellings. The pages are searchable, though accuracy is limited by textual inconsistencies. As in version 1, users access individual records from a homepage containing an itemized list of records for each series. As Figure 2 shows, users click the desired record to open the record page in a new browser tab. Rather than a simple image, the version 2 record page features a light box gallery including all images associated with each record, as well as next and previous navigation links allowing users to read the



Version 3 is an interpretive version. It contains digital reproductions of the original documents paired with corrected and edited transcriptions. These materials are augmented with supporting information, such as notes and references gleaned from secondary sources, maps, and personal biographies. Version 3 is thoroughly cross-referenced, with links to support easy navigation across diaries and letters. Thematic tagging enables users to isolate records dealing with specific topics. Only those records deemed most interesting were included in this presentation. The version 3 homepage offers more in-depth contextual information than either of the other presentations, with a narrative introduction rather than a finding aid structure. Records pages in version 3 are similar to those in version 2, with next and previous navigation and the light box gallery. As shown in Figure 3, however, Version 3 record pages contain more metadata fields, a block of introductory text, footnotes featuring information not present in the record itself, and a corrected and edited transcript. Version 3 also provides additional resources stored on separate subpages within the version 3 directory, such as a directory of people mentioned in the collection with biographies; a list of places mentioned in the collection with links to coordinates on a customized map; and a page of selected “essential” records featuring some of the most interesting records in the collection.



## June 31 1854 – Martha Haun to James Haun

<Previous Next>

*In this letter, much delayed due to a change in the postal schedule, Martha Haun discusses tentative plans to travel to California either with some neighbors or with her brother-in-law Dave, and debates whether to bring slaves or leave them behind in Kentucky.*

My Sweet Husband,

I received your kind of precious letter on Wednesday, last the 28th of June. It was dated the 28th of May<sup>2</sup>.

June the 31st 1854<sup>1</sup>

James Haun recorded the receipt of this letter in his diary [September 3 1854](#).

*Figure 3.* An example of the version 3 presentation with sample notes feature. Reprinted with permission of the Plumas County Museum Association.

**Platform.** The Haun Collection test site is a WordPress website built using a modified version of the Twenty Thirteen template site theme. The site has its own banner. Fonts and coloration belong to the original theme. With the exception of the map, most features are powered by existing WordPress plugins (see Table 1).

Table 1

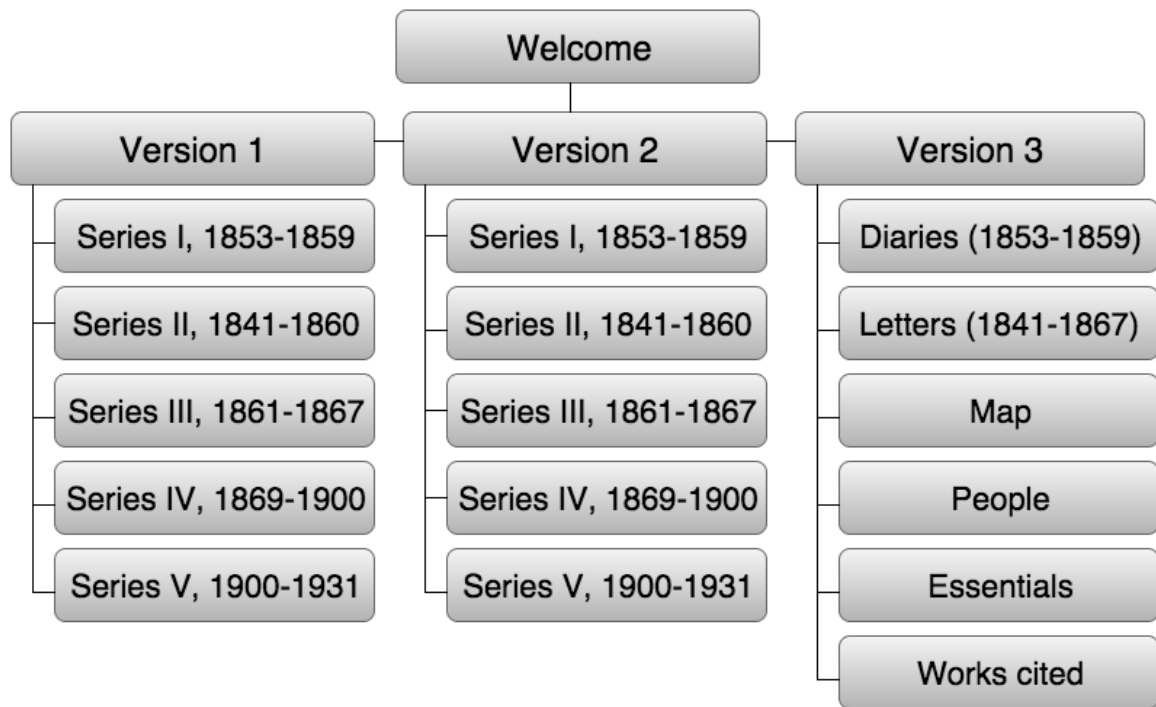
*WordPress plugins*

Plugin name	Functionality
Add from Server	Allows for bulk uploads to the Media Manager using FTP (essential in a site where all images exceed the default maximum upload size)
Add Meta Tags	A search optimization tool which allows for backend Dublin Core metadata, among other tools
Collapse-O-Matic	Generates the expandable lists featured on each of the version homepages
Columns	Allows for the 2-column presentation of the diary entries in version 2
Disable Comments	Allows for selective comment suppression; comments are currently disabled throughout the site
Exclude Pages from Navigation	Prevents all 1,120 pages from appearing in the top nav bar, rendering the site unusable
Footnotes	Powers the footnotes that appear in version 3
Gallery Carousel Without JetPack	Generates the light box galleries in versions 2 and 3
Insert Headers and Footers	Allows for new elements in the header and footer sections of the site
Tag Pages	Powers the subject and thematic tagging in version 3

The other features—including the previous and next navigation in versions 2 and 3, metadata, and summary boxes—were generated using simple html.

In evaluating potential technologies, it was important to select a tool affordable enough for small archival budgets with high design standards to meet users' expectations for online interfaces. In addition, the technology must be approachable enough to be administered without professional programmer support. Several platform alternatives were considered, including Dreamweaver and Squarespace. Sample pages were generated using each tool, but only WordPress combined effortless, modern graphic design with the flexibility to present the different version, and the ability to easily create internal navigation between versions and records.

**Site layout.** As Figure 4 shows, the site structure consists of a welcome screen and homepages for each of the three versions, with individual record pages and other resources nested beneath. Basic site architecture was put in place in January 2014. Materials were added on a rolling basis beginning in March with the version 1, Series 1 photographs and concluding in October with the final touches to the version 3 maps, people lists, and subject tagging. The site remained under password protection throughout its development.

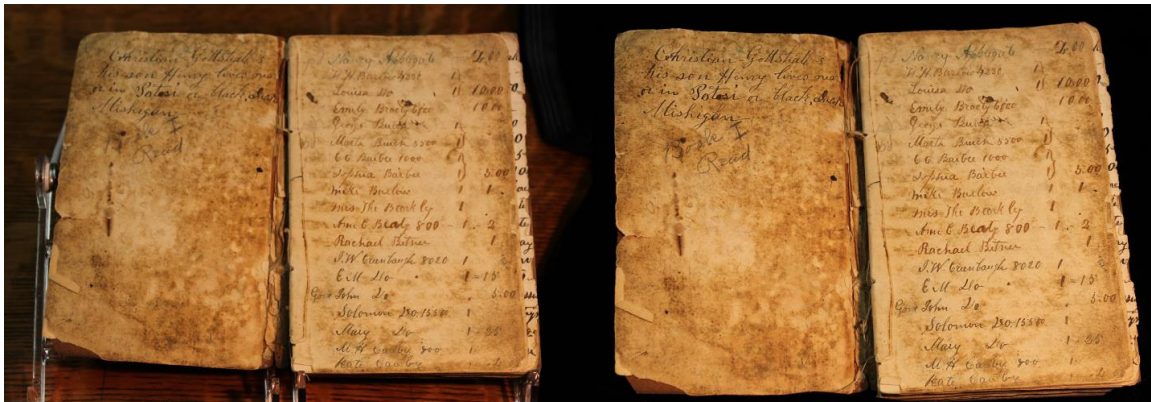


*Figure 4.* A diagram of site layout.

Volunteer testers recruited from among friends and family reviewed early versions of the site. Errors in execution were identified through this process and corrected prior to the general release.

**Digitization and content preparation.** The first stage of the project focused on digitizing the collection in preparation for test site creation. Digitization took place primarily during summer 2013, beginning with the letters that make up Series 2 - 5. Documents were laid flat on the bed of an Epson V330 photo scanner for reproduction. Copies were saved separately in both TIFF and PDF formats.

Due to the delicate handmade bookbindings, the diaries that make up Series 1 could not be laid flat for reproduction. Instead, they were propped on adjustable bookstands under a covering of black velvet and photographed at an angle to approximate a straight-on perspective as shown in Figure 5. Images were taken of each recto and verso individually as well as each complete page opening. Ultimately, only the complete page opening shots were used on the test site. A Canon Rebel EOS T3 was used for the photography. All records were handled with care while wearing protective gloves. Total time for photos and scans amounted to approximately 35 to 40 hours.



*Figure 5.* An example of the photography setup. Left: a test shot of the first volume of James Haun's diary without backdrop, bookstands visible. Right: the final image with cloth background in place.

Images from Series II - V were subsequently arranged in chronological order and cropped to minimize white space using the open-source photo editing software, Gimp. In some cases, Series II - V originals were too large for the scanner bed. In these instances,

images were scanned in stages, then cropped and spliced back together. Many letters in the collection were written on folded sheets of post-sized paper, meant to be opened like a book, with page 1 on the front, pages 2 - 3 within, and page 4 on the back. Though these documents might have been reproduced in their entirety as full images, pages 1 and 4 in one image, pages 2 and 3 in another, they have been treated as separate images on the test site in order to preserve chronology in the interface and to maintain a consistent treatment for letters longer than four pages.

Because bulk-uploading functionality native to the WordPress platform does not support a folder and subfolder arrangement, image files from Series II - V were labeled according to series, number, date, and page number. For example, in Series III, the ninth item in chronological order is a letter dated October 5, 1863. The first page of that letter is saved under the file name "siii\_009\_October 5 1863 p 1." In this way, images automatically appear in the appropriate order without relying on subfolders.

Series I images were processed more quickly and with less effort. Some images were cropped to reduce white space, but most were useable in their original form. The diaries were photographed in chronological order. Images were similarly retitled for upload, according to writer, book number, and date. For example, the 46<sup>th</sup> image in chronological order from James Haun's first journal is labeled "jh\_bk1\_046\_January 20 1854 - January 25 1854."

The next major step in preparing site content was generating the necessary transcripts. The majority of Series I had been previously transcribed by museum staff in the 1970s. The transcription was Xeroxed and given to a freelance typist for rekeying into an editable Word document prior to this project. During fall and winter 2013, Series



II - V were transcribed manually. Select portions of Series 1 had been overlooked or misplaced during the earlier transcription effort, including the accounts book, Book IV, and John Haun's pencil diary, apparently a draft of his pen and ink version, Book VI. These were also retyped. Two copies of each completed transcript were saved. One became the copy of record for version 2; the other formed the basis for the edited text in version 3.

The version 3 editing process began with straightforward and arguably more justifiable changes, including imposing standardized conventions for dates and currency, consistent spelling of names, standard spacing, and correct spelling and punctuation. In some cases, incorrect spelling was retained, for example, "diggins" instead of "diggings," "the boys was hauling" instead of "the boys hauled." Still, many of the more charming eccentricities were eliminated, especially from James Haun's diary—"to kt" became "tonight," "kiota" became "coyote." While this makes the material far easier for a modern reader to digest at a glance, it also strips the writing of much of its character. It is easy to imagine James Haun's speaking voice from his spelling.

With these more cursory elements in place, a second sweep addressed confused and unclear passages in the text, added footnotes, and determined which records to include in the final presentation. This process changed James and Martha Haun's writing most substantially. Martha's histrionic flights and repeated professions of love and anxiety are at times almost impossible to follow in the original and were substantially edited for clarity in version 3. Though no outright cuts were made, many sentences were restructured. This was occasionally, though less frequently, also the case in James Haun's diary.

## **Outreach to Potential User Groups**

This study was reviewed and approved by the San José State University Institutional Review Board (IRB) before research participants were recruited. Individuals and organizations associated with the previously identified user groups were contacted in the months leading up to the site's release and throughout the testing period.

The first audience identified consisted of residents of Plumas County including both existing patrons of the Plumas County Museum and potential new users in the area. The Plumas County Museum agreed to direct interested researchers to the site as a resource. The project was highlighted in a half-page article appearing in both the *Feather River Bulletin* and *Indian Valley Record*, weekly local newspapers published by the Feather River Publishing group of Quincy on Wednesday November 13. On the same day, links to the collection and survey were posted to the Quincy community Facebook page, "Quincy Peeps," a private group with over 2,000 members (about 40 percent of the town's population) shown in Figure 6. The project and site were featured in the Plumas County Museum quarterly member newsletter mailed in hardcopy on November 21, 2014. A pdf version was later placed on the museum's website.

The second potential audience included Plumas County students and educators, especially those focused on California or American history. Outreach began in the months leading up to the public release of the site, so that teachers could have time to integrate the project into lesson plans. Instructors from the Quincy Jr/Sr High School and Feather River College in Plumas County were contacted via email and invited to participate in the project either individually or with their classes.

A third potential usership made up of academic researchers specializing in topics such as history of the west, California history, and the gold rush was also projected. A list of academics with relevant expertise and interests was developed through online research into top PhD programs in these disciplines. After the site's release, researchers were contacted individually via email and invited to participate. Email outreach continued throughout the testing period. Over 45 invitations were sent to scholars specializing in topics relevant to the collection

Lastly, it was hypothesized that other, smaller potential audiences might include professional genealogical researchers, local history groups from Scott County, Kentucky, and California history enthusiasts more generally. Outreach to these disparate groups began at the institutional level before the site's official release by contacting related institutions and member organizations. In the area of California history, these groups included the California Historical Society, the San Francisco Historical Society (which has some interest in broader statewide history), the Conference of California Historical Societies, California Genealogical Society, and the Oakland Museum of California. In Kentucky, the Kentucky Historical Society, Georgetown Museum, Georgetown High School were contacted. These organizations were invited to participate in the study and encouraged to promote the research to any members or patrons they felt might take a special interest in the project. This generated a short list of interested individuals waiting to be notified when the site was officially released on October 31, 2014. The list included institutional staff, armature historians and a descendant of the Haun family. A small subset of professional genealogical researchers and bloggers writing on related topics were contacted individually by email. The site was not promoted extensively to

general audiences but was mentioned repeatedly on Twitter during the testing period as shown in Figure 6. Finally, the site was crawled and indexed by Google, allowing for serendipitous discovery.

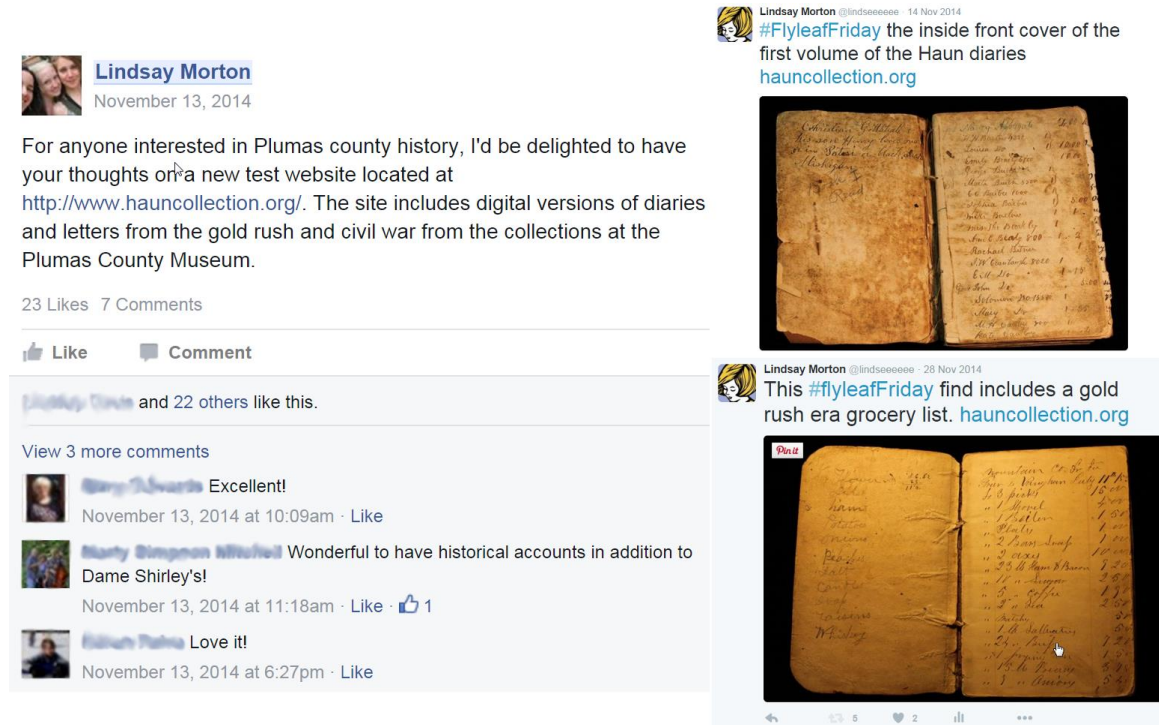


Figure 6. Examples of social media outreach via Facebook (left) and Twitter (right).

Outgoing communications through all channels encouraged users to access and explore the test site and asked them to submit feedback by completing a short exit survey. Direct links to both the test site and the survey were provided in the various communications. The site's homepage also explained the study's purpose and provided access to the survey for the convenience of the invited testers and the benefit of any serendipitous browsers.

### Data Collection

A five-week open testing period from October 31, 2014, to December 5, 2014, generated 30 results from a voluntary exit survey administered through SurveyMonkey

(surveymonkey.com) and web traffic patterns captured in Google Analytics. The IP tracking feature native to SurveyMonkey was disabled for complete anonymity. IP addresses captured by Google Analytics were excised from results prior to analysis.

The survey instrument included 12 questions designed to assess the collection's usefulness, gauge users' reactions to the three versions presented, and gather demographic information and details on users' research background. Specifically, questions 1 and 2 focused upon the collection itself, asking the tester to rate the material's usefulness and the likelihood of their consulting a similar site in the future. The intent of these paired inquiries was twofold: 1) to quantify the value and appropriateness of the collection for each user group consulted and 2) to test whether users' degree of interest in the content influenced presentation preference. Response choices were offered along a standard 5-point scale ranging from 1 to 5, with 1 being the best.

The next set of questions centered on access and personal preference. Users were asked to first confirm which of the three presentation versions they viewed, specify which single version they found most useful, and, finally, to rearrange the three options in order of personal preference. Question 3 was purely practical. Users who viewed only two of the three presentation options had the potential to skew results. It was therefore important to account for this during analysis. Questions 4 and 5 were intended to work in concert, testing users' reactions when forced to make just one selection versus prioritizing all three (1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup>), and exploring the difference between usefulness and personal preference. As shown in Figure 7, images of the three versions were provided alongside version-specific questions in order to help testers recall which version included which features.



**3. Access to versions.**

	Version 1	Version 2	Version 3
Which of the three views of the collection did you see? (Mark all that apply)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**4. Version usefulness.**

	Version 1	Version 2	Version 3
Which of the three views of the collection did you find most useful to you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**5. Rank each of the three styles in order of your personal preferences**

<input type="text" value="Version 1"/>
<input type="text" value="Version 2"/>
<input type="text" value="Version 3"/>

Figure 7. Sample Survey Screen.

Questions 6 and 7 requested users to comment upon the specific features they liked and disliked, while question 8 offered the opportunity for any additional unstructured comments. All three questions in this section were optional and were intended to elicit testers' detailed and unguided thoughts and reactions to the site.

The final section asked users to provide general information on their personal experience, including research role, previous interactions with primary source documents, and formal education level. This portion of the survey aimed at identifying similarities and differences among users in order to interpret the results according to research role and experience. Participants were offered a choice of five categories to describe

themselves: scholar, student, teacher, information/avocational researcher, or other (with a blank space to provide details). In a larger study, further differentiating between elementary, secondary, and post-secondary teachers may be informative. (See Appendix A for the complete survey form with available response options.)

Prior to administration, the questions were vetted and revised at the recommendation of the thesis review committee which included specialists in digitization, online and new media curation, history, and archival and records management. The survey interface was tested for functionality by the researcher.

Data on traffic patterns and usage was gathered using Google Analytics and filtered to exclude the researcher's own visits to the site.<sup>5</sup> This data provides a more objective perspective on the amount of traffic the site received, how much of the collection was viewed, and overall length of visits.

### **Limitations and Delimitations**

In addition to the normal limitations intrinsic to survey research and self-reporting, the current study may have been influenced by participant recruitment methods. Participants were actively recruited from among previously identified groups of potential users with the aim of securing representation from all relevant user populations. This approach was essential to the completion of the study but influenced analytics results by driving traffic to the site. Receiving an invitation to review the site

---

<sup>5</sup> Traffic was filtered in three ways. Between the date of the site launch on October 31 and November 15, I was traveling on business and accessed the site from Canada, the UK and Portugal. As all foreign traffic appears to be my own, this was easy to eliminate. Beginning November 25, IP address filters blocking my home and office were put in place, preventing the site from tracking me going forward. Between these two periods were 9 days during which my own traffic had to be eliminated manually.

may potentially have influenced user behavior, affecting the number of pages visited, types of activities performed, or time spent exploring. The risk of unrealistically skewing analytics results is somewhat mitigated given media coverage similar to the outreach used to recruit study participants might reasonably accompany the release of a real online archive.

The survey questions were not formally pretested, as the relevant population was limited and individuals could not participate in both a pretest and the study. Feedback on questions, methodology, and execution plan was obtained from a committee of experts. The survey interface was tested from the user perspective to confirm that all questions functioned as expected.

Due to these factors, the results were preliminary rather than definitive, though useful conclusions can be drawn (see Chapters 4 and 5). Following refinements to the site based on the results of this survey, future research may explore site activity over a longer period of time, as well as usage based on serendipitous discovery rather than invited reviews. Finally, this study focuses primarily on utility. Future researchers may wish to explore additional factors contributing to user preference, for example, user experience (UX) design.



## Chapter 4: Results and Discussion

The initial results reported here cover a period of 36 calendar days from Friday, October 31 through Friday, December 5, 2014. The open testing period resulted in 25 completed surveys, five partial surveys, and Google Analytics data from 144 new users and 183 total users over 218 sessions.

Survey responders self-identified into several preselected groups outlined in Table 2: informational/avocational researchers (23%), students (30.8%), teachers (11.5%), scholars (30.8%), and one “other”—a genealogist with a degree in library science. Several respondents seemed confused by the “informational/avocational” terminology, describing themselves as “other—person interested in local history” or “other—nonprofessional reader interested in Plumas County history.” These were re-classified manually and considered within the “informational/avocational” group.

Table 2

### *User group representation*

	Percentage	Total responses
Scholar	30.8%	8
Student	30.8%	8
Teacher	11.5%	3
Informational/avocational	23.1%	6
Other	03.8%	1

Survey questions in the first section ask visitors to assess the usefulness of the collection overall. Results were largely positive, with 24 of 30 responders (80%) rating the collection as “Extremely useful” or “Very useful” and 23 of 30 responders (76%) describing themselves as “Extremely likely” or “Likely” to use a similar site again in the future (see Tables 3 and 4).

Table 3

*User perceptions of overall usefulness*

	Extremely useful	Very useful	Useful	Somewhat useful	Not useful	Average rating
Percentage	36.67%	43.33%	16.67%	3.33%	0	4.13 of 5
Total responses	11	13	5	1	0	

Table 4

*Likelihood of visiting similar sites in future*

	Extremely likely	Likely	Somewhat likely	Somewhat unlikely	Very unlikely	Average
Percentage	40%	36.67%	20%	0	3.33%	4.10 of 5
Total Responses	12	11	6	0	1	

When analyzed within the context of user group (see Table 5), these two metrics continued to track closely with one another. Students and those who did not complete the demographic portion of the survey expressed the least enthusiasm for the resource, though reactions could not be described as negative.

Table 5

*Comparison between assessment of usefulness and interest in similar sites by user group*

Ranking	1		2		3		4		5	
<b>Informational</b>										
Interest	3	50%	2	40%	1	10%				
Usefulness	4	66%	2	33%						
<b>Other</b>										
Interest			1	100%						
Usefulness			1	100%						
<b>Scholar</b>										
Interest	4	50%	3	37.5%	1	12.5%				
Usefulness	3	37.5%	4	50%	1	12.5%				
<b>Student</b>										
Interest	2	25%	2	25%	3	37.5%			1	12.5%
Usefulness	2	28.5%	2	28.5%	2	28.5%	1	14%		
<b>Teacher</b>										
Interest	1	33%	2	66%						
Usefulness	1	33%	2	66%						
<b>No response</b>										
Interest	1	33%	1	33%	1	33%				
Usefulness	1	24%	2	50%	1	25%				

Twelve of 26 survey participants (46%) ranked version 3 most useful, followed by version 2 (34%), and version 1 (20%). When further subdivided by user group as shown in Table 6, results indicate that the genealogical researcher found version 1 most useful, while scholars and students were divided between versions 2 and 3, and the other groups gravitated more definitively toward version 3.

Table 6

*Usefulness assessment by user group*

	Ranked version 1 most useful	Ranked version 2 most useful	Ranked version 3 most useful
Informational	2	1	3
Scholar	1	4	3
Student	1	4	3
Teacher	0	1	2
Other	1	0	0
Total	5	9	12

When users were asked to rank each version in order of personal preference a discrepancy emerged (see Table 7). First choice selections were almost equally distributed across all three versions, while version 3 was most often selected as 3<sup>rd</sup> or last choice.

Table 7

*Version preference*

	1 <sup>st</sup> choice	2 <sup>nd</sup> choice	3 <sup>rd</sup> choice
Version 1	34.62%	30.77%	34.62%
	9	8	9
Version 2	30.77%	50%	19.23%
	8	13	5
Version 3	34.62%	19.23%	46.15%
	9	5	12

To speculate, users may prefer one version to another for reasons other than utility (see Limitations and Delimitations discussion in Chapter 3). Viewed within the context of user group (see Table 8), the results suggest that versions 2 and 3 were most popular among students and teachers while scholarly and informational preferences were more evenly distributed across versions.

Table 8

*Version preference by user group*

	Prefers version 1	Prefers version 2	Prefers version 3
Informational	3	1	2
Scholar	3	3	2
Student	1	4	3
Teacher	1	0	2
Other	1	0	0
Total	9	8	9

When users reported on the frequency with which they consult primary resources a small majority consisting of 14 of the 26 respondents (53%) indicated that they access primary source material six times per year or more (see Table 9).

Table 9

*Frequency of consulting primary source material (per year)*

Users	Less than 1	1-2	3-6	6 or more
Informational	3	2	1	0
Scholar	1	0	0	7
Student	1	2	1	4
Teacher	0	0	0	3
Other	0	1	0	0
Total	5	5	2	14

When users evaluated the positive and negative aspects of version 1, they complimented image clarity and quality, as shown in Table 11. Negative comments, shown in Table 10, focused on the user interface and navigation, including difficulty in retrieving material quickly, the amount of clicking required for accessing the images, the lack of “page flipping” or previous and next navigation options, and the tedium of having to return to the series home page in order to access each new image. Lack of detail was also a frequently cited issue. Two respondents noted difficulty in reading the handwriting in the records, and one, who reported never having worked with primary source material

before, noted that the poor spelling was a distraction. Another, who reported accessing archival material six or more times per year, complained that the descriptor “series” was difficult to understand.

Table 10

*Negative written feedback by category*

	Version 1	Version 2	Version 3
User interface	9	3	6
Context	1	1	2
Content	6	1	4
Organization	0	0	0
Digital quality	0	0	0
Other	1	0	0
Total	17	5	12

Table 11

*Positive written feedback by category*

	Version 1	Version 2	Version 3
User interface	4	5	2
Context	3	0	3
Content	4	14	14
Organization	3	3	0
Digital quality	6	0	0
Other	1	1	0
No complaint <sup>6</sup>	0	6	5
Total	21	29	24

Version 2 feedback included positive comments on the content (especially the transcriptions) and user interface. For example, one respondent wrote, “The transcriptions are very useful. The organizational scheme makes searching for content easier if you know who/when you are looking for. Navigating within digitized content is

---

<sup>6</sup> The “No complaint” category has been included in cases where the respondent wrote that they had no negative comments to offer, implying that the version they reviewed did not require critique or improvement. This was expressed in terms such as “all good” or “No dislikes.”

easier than in Version I [*sic*].” The respondent further stated, “I love the verbatim transcriptions combined with the clear images and the easy navigation from page to page. I can scan the text and then refer to the image when needed or pull up the large resolution image for a closer look.” Negative comments included difficulty in reading the transcriptions, though it is unclear whether this may be due to the arrangement on the page or the challenging spelling and phrasing. Accessing the images from the homepage was still felt to require too many separate clicks. One scholar described version 2 as “difficult to navigate” and “harder to understand...in its entirety.” Another commentator, who described his/herself as a teacher with a PhD, felt that “version 2 didn’t have enough descriptive text—I wasn’t sure what I was looking at/its significance.” Six respondents stated that they had no negative comments, making such remarks as “all good,” “no dislikes” or “looks good.”

Version 3 responses were more varied. Users reported a high degree of satisfaction with the content, including the transcriptions and bibliographical information, “useful accompaniments,” top of page descriptors, and maps. Two users on the other hand, both scholars, could find nothing positive to say about version 3, noting “did not care for typed version” and simply “did not like.” Negative comments were diffuse. Two users reported difficulty navigating and understanding the content or its chronology. One found “the lack of organization by category” to be a detractor. Since version 3 is the only version available with thematic tagging, this is presumably a user interface issue rather than a content issue. One felt that version 3 contained “too much text.” Two more did not like the edited transcripts. Five responders reported having no complaints about this version.

Users also had the opportunity to make any additional comments. The more substantive and detailed responses included the following:

- “After a quick survey I'm amazed at the work involved. This takes time and money. Most library and archival institutions I know are poorly funded and at best can afford only to develop finding aids. I would love to see every collection processed to this stage, but that's a fantasy I'm sure.”
- “Although version three seems to be for the mass market, it may be too busy for a serious scholar only interested in the family's writings. Photos, reference maps, etc. helps peak interest of broader audience.”
- “Digital archives offer a lot of possible additions, but really, we're after the meat. Quick easy access to the digitized material itself. One should be able to flip through and know where you are.”
- “Awaiting Great War Era dairies and letters. Also, zoom on letters is extremely close. Perhaps a way to zoom in and out?”
- “This is an amazing find. I have a historical family connection to the area. My great grandfather was a contemporary of David Haun and the Pratt family.”
- “I very much appreciate the abstracts of the texts which allow one to search for terms which are not explicitly stated in the documents. For example, out of curiosity, I search for "slave" and was able to find a letter describing the situation of several of the family's slaves even though the word "slave" never appeared in the letter itself. Also, the tags are useful as well. How about a family tree and/or glossary of names of people mentioned in the texts? Obviously, someone doing



full research with the letters is going to know that background, but students and other general users might benefit from a Who's Who, if you will.”

- “Involving images on every page would probably make it more engaging—and replacing generic text with more descriptive titles would also help me understand what I was looking at.”

(See Appendix B for complete written responses.)

Google Analytics results indicate that more users visited the collection than completed the survey by a factor of almost 5 to 1 for first time users or 6 to 1 for new and returning users combined. Peak activity on the site correlates closely with public outreach and media coverage as shown in Figure 8, with clear spikes following media coverage of the project, including a November 12 newspaper article and Facebook posting and a particularly impactful November 22 Twitter posting.

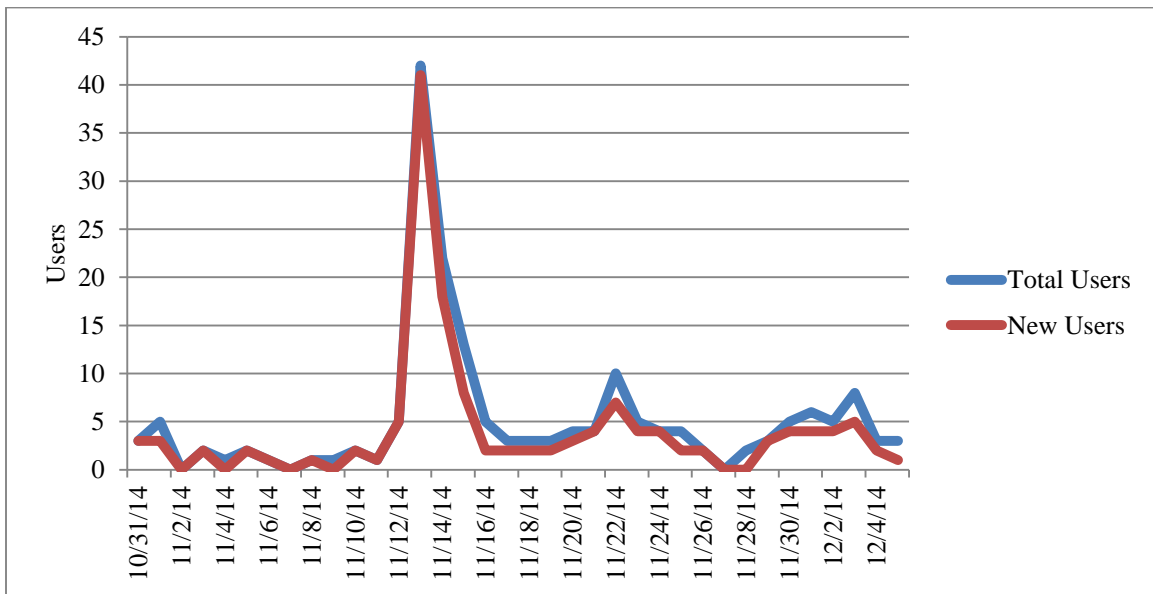


Figure 8. Site visits over time.

A majority of visitors (68%) originated from California. However, the site also attracted viewers from 14 other states including Arizona, Oregon, Colorado, Indiana,

Illinois, Massachusetts, Ohio, North Carolina, Kentucky, Nevada, New York, Pennsylvania, Washington, and Florida. The majority of California-based traffic originated from Plumas County, although the site did attract some attention from more distant parts of the state including the Central Valley and greater Bay Area, along with slight traffic from southern California.

In total, the site attracted 1,703 page views and 1,331 unique page views across 229 different pages. An average user viewed between six and seven pages during a single visit. These numbers do not include media files, which are shared across all three versions. Because version 1 consists almost exclusively of media files, metrics are not representative of actual activity for that treatment. Versions 2 and 3, however, may reasonably be compared against one another. As Table 12 shows, version 3 was accessed nearly twice as often as version 2, and users explored a larger portion of this version. Since the web pages showing diary entries in version 3 are presented in increments of one month rather than individual page openings, users visiting version 3 records actually access the equivalent of 3-6 records pages in version 2, further increasing the amount of material viewed in this format.

Table 12

*Site access by version*

	Count of Unique URLs Accessed	Page Views	Unique Page Views
Version 1	6	213	173
Version 2	86	420	286
Version 3	137	783	671
Homepage	1	237	155

The hierarchical structure of versions 1 and 2 is evident in usage patterns. In examining a list of the top 30 most accessed URLs within the site (see Table 13), it

becomes clear that many are gateway or directory pages which users must pass through in order to reach the more content rich records pages. The first several pages of the first series presented across each version also appear repeatedly in the top results. Page view numbers dwindle deeper into the site, as users drop off or disperse into the wider collection. Interestingly, version 3 records pages appear to suffer less dramatically from this drop off, perhaps due to the elimination of the series-level hierarchy in this instance, or as a product of the “essentials” list featuring links to 16 of the most “interesting” records. In total, pages belonging to version 3 account for more than half of the most viewed site pages, demonstrating that, while version 3 received mixed ratings and reviews, it also drew the most attention and deepest engagement from users

Table 13

*Most accessed site pages in order of page views*

Page Title	Page Views	Unique Page Views
The Haun collection - main homepage	237	155
Version 1 homepage	99	82
Version 3 diaries	97	60
Version 2, series I, homepage	92	34
Version 1, series I, homepage	74	56
Version 2 homepage	63	52
Version 3 homepage	50	44
Version 3 letters	30	17
Take our survey	27	25
Version 3 map	25	24
Version 3 people list	22	19
Version 3, February 1853	20	18
Version 2, March 18 – 21 1853	15	12
Version 1, page 1: accounts	15	14
Version 3, accounts	14	13
Version 1, series II, homepage	13	12
Version 2, series II, homepage	13	8
Version 3, March 1853	13	11
Version 3, February 1855	12	10
Version 3, January 1854	12	11
Version 3, March 1854	12	9
Version 1, series III, homepage	11	8
Version 3, August 1853	11	9
Version 3, July 1853	11	8
Version 3, June 1853	11	9
Version 3, November 1853	11	9
Version 1, series V, homepage	10	10
Version 2, September 1853	10	9
Version 3, essentials	10	9
Version 2, series IV, homepage	9	8

The results suggest several opportunities for improvement to the study design and administration. The language used to identify user groups within the survey was confusing for some respondents and should be revised for clarity. For example, as

previously discussed, some users misunderstood the terms “informational/avocational” and “student.” They marked themselves as “other” and provided written explanations of their research roles and motivations that were manually reassigned to the most appropriate category before analysis. Likewise, question 5, which asks users to arrange the versions in order of preference, proved unintuitive. A refined survey would simply ask users to identify their most preferred and least preferred versions; the second choice could be easily inferred when analyzing the results. The site traffic data for version 1 in particular could be enhanced using Google Analytics event tracking.

Overall, however, the outcomes of the testing period are encouraging and constructive. Clearly, the collection itself has relevance for a broad spectrum of users, including individuals with various experience levels, ages, and research goals, both within Plumas County, and beyond. Responses indicate that version 3 was slightly more useful overall (see Table 6), though this did not necessarily correlate with personal preference (see Table 8). When assessed through the filter of user group, scholars preferred each version at nearly the same rate, while students preferred versions 2 and 3, and teachers ranked version 3 most highly. The implications and potential applications of these results are discussed more fully in Chapter 5.

## Chapter 5: Conclusion

In evaluating the Haun Collection test results, it is helpful to consider outcomes within the context of the organization's overall programing. Many website administrators would view such low numbers—144 new users, 183 total users, and only 20% of available pages accessed—as disappointing. However, more users accessed the Haun collection during the five-week testing period than accessed all of the records in the museum's physical reading room combined over the previous three years. Usership for November 2014 jumped to 36 times the normal level. (Of course, that kind of growth is easy where numbers are small to begin with. Sustained growth over time is a different matter.) It is important to remember that archival reference services measure success on a different scale from most mainstream online entrepreneurs. This is a marketplace in which the Huntington Library, a premiere institution housing six million manuscripts, boasts of “1,700 scholars every year” in its marketing materials (Huntington Library, 2015). Thus, while numbers from the first month of the collection's run are not to web scale, activity in the site demonstrates a potential audience for resources like the Haun Collection that substantially exceeds that of the physical reading room, as well as the physical boundaries of Plumas County. In discussing a special digital exhibition of the University of California Berkeley's Institute for Research on Labor and Employment Library, Huwe asserted, “digital collections gain vast new readerships when they appear online in structured and searchable formats” (2009, p. 15). The present study demonstrates that this is potentially true, not only for major collections at large, well-funded, internationally recognized institutions, but also for the hidden collections housed at small or niche institutions. Because access and site traffic statistics are most often

private and internal, it is difficult to accurately assess how the Haun Collection site traffic compares to that of comparable resources.

In addition to extending the collection's reach and helping to preserve the records, bringing the collection online has unexpectedly uncovered linked resources and related research projects that might otherwise not have become known. For example, one tester has been investigating the Haun family and its connections in the decades before they migrated to California. In another case, a link has been uncovered between the Haun Collection and a related manuscript penned by James Haun's sister-in-law and cousin, Catherine, recounting her journey overland to California in 1849.

The present study contributes to a more holistic understanding of user needs across a broad spectrum of archival users. On the question of user identity and presentation preference, the evidence of the Haun Collection test shows that users found a highly processed, highly contextualized presentation method more useful than the alternatives by a slight margin. When the test population's preferences were broken out according to user group, the assessments of the scholarly community and the single professional researcher tended to diverge from the reactions of other testers, though no group was uniform in its response. Among scholarly and professional users, versions 1 and 2 were preferred. Written comments suggest that these groups viewed the additional features in version 3 as extraneous or detracting from the main point of the records. We can speculate that these users' deep knowledge of the time periods covered and previous research experience may have obviated the need for contextualization. Simple familiarity may also have had a role to play. Once a user becomes comfortable with a certain way of working, the demands of learning a new system can outweigh the benefits, especially if

improvements are incremental. Finally, it is worth considering the less practical, more elusive rewards of scholarship. Exploring history through primary sources has been compared to solving a puzzle or unraveling a mystery. Insight and enlightenment are core to the experience, and the pleasure of discovery has no doubt influenced many scholars in their decision to pursue a career in research. In demanding less of users, a high degree of processing may reduce engagement and perhaps detract from the rewarding feelings associated with investigation.

Students and teachers preferred versions 2 and 3 overall. Written comments suggest that the additional information provided helped users to connect with the material, highlighting its relevance to them. At the same time, the volume and breadth of the collection seemed to overwhelm some users. Comments on the large amount of text and the need for more images relative to text suggest that greater selectivity would be acceptable, perhaps even appealing, to users in these groups. Informational and avocational researchers did not show a clear preference.

Results suggest that the optimal presentation for the constituency of the Plumas County Museum lies somewhere between versions 2 and 3—combining the authenticity and accuracy of version 2 with selected elements of the enhanced content and design clarity in version 3. A revision should also expand modular access, increasing users' ability to approach the records outside the hierarchical top-down navigation in order to reach the material that interests them more quickly and directly. The site's next iteration should explore the addition of a progress bar or timeline along the top on the screen, designed to help orient users within the collection, an element which is currently missing from all versions. Finally, while the present study was not primarily intended to evaluate



interface design, it is clear that features like tagging must be emphasized further in order to promote use.

Although the current study differs from related research in key ways, including the focus on smaller organizations and the comparison of multiple presentation models across a spectrum of audiences, certain parallels can be drawn. The response from the scholarly community tracks closely with results from other user studies. For example, West and Fesenko (2009) found that scholars preferred digital archival organization to “replicate the physical arrangement of the manuscript collections as closely as possible” (p. 14), as in versions 1 and 2 of the Haun Collection test. Scholarly users strongly favored the practice of digitizing complete collections, including items that might appear insignificant, as opposed to the selective digitization of only the most important records. Again, this finding aligns neatly with scholarly testers’ preference for versions 1 and 2 over version 3 in the Haun Collection test. Finally, West and Fesenko’s scholarly testers paralleled the Haun collection testers in their ambivalence toward enhanced features such as tagging.

Similar parallels exist between the responses of the high school students who tested the Haun Collection and research conducted with college undergraduate archival users. Archer (2009) found that undergraduates needed substantive guidance to identify, locate, and access primary source documents. High school students’ preference for versions 2 and 3 of the Haun Collection suggests a similar appreciation for a higher degree guidance or mediation.

The discrepancy between the scholarly participants’ responses and those of other user groups are indicative of an “inherent divide between meeting the needs of

researchers and those of the general public” (Roth-Katz, 2012, p. 124). This divide has been explicitly identified in the literature as well as intrinsically expressed in research designs that focus on particular groups to the exclusion of others, tacitly acknowledging these divisions (Allison-Bunnell, Yakel, & Hauck, 2011; Barth, 1997; Duff & Johnson, 2013; Goggin, 1986; McKay, 2002; Stieg Dalton & Charnigo, 2004; Tibbo, 2003; Yakel, E., & Bost, 1994). This also speaks to the need articulated by Peacock and Brownbill (2007) for more advanced market segmentation and a strategic, integrated approach to understanding “usability, motivation and demand.”

It is still too soon to gauge the long-term potential for digital assets like the Haun Collection, particularly versions 2 and 3. While many articles discuss the design and construction of online archives, few follow-up reports have appeared, limiting the information available for comparison. The prototype Polar Bear Expedition Digital Collection site discussed in the literature review was “more and more difficult to keep up-to-date” after the “project investigators moved on to other activities.” The University of Michigan rebuilt the site to make it “more easy for researchers to use, and more easy for staff to update” (Bentley Historical Library, 2015). Other sites like the Bancroft’s “The 1906 San Francisco Earthquake and Fire” have remained live and unchanged for years. Among the archival sites that remain both active and dynamic, most usage information is considered proprietary. For example, the California Digital Library offers traffic statistics to its contributors at an institutional level but does not release details publicly or offer comparative data (California Digital Library, 2013). Perhaps the closest parallel to the highly interpretive features in version 3 can be found among art archival websites, where curatorial sensibilities make augmentation and contextualization more common.

Future researchers may wish to pursue a meta-analysis of available site traffic data in order to gain insight from this parallel field.

In considering a digital strategy, archive administrators must make calculated business decisions. Whether the gains in exposure and preservation will justify the investment in time and resources at an individual organizational level will vary from collection to collection and institution to institution. Similarly, in assessing the suitability of various digital presentation models, organizations must prioritize user needs according to several factors, including needs and experience levels of the usership, the relative volume of users from each group (both current user population and the larger potential usership), organizational mission and goals, and desired outcomes. If, for example, the institutional mission centers on supporting scholarly research, a creative archiving approach may not be appropriate. However, if engaging the community in the past or passing heritage on to the next generation are key goals, this model may serve the needs of the organization most effectively despite the challenges inherent in initial setup.

Ultimately, digitizing small repository collections is only a first step. Publicly available online records—and data *about* records—would open the possibility for deep cross reference, for data mining, for a computational approach to history that is impossible when information is scattered across innumerable individual repositories. It could also lead to more informed investment in culture and grant allocation based on broad, relativistic value assessments. For small, local museums and archives, there exists enormous untapped potential for new audiences, for unforeseen connections, for a network of cumulative knowledge that would enable us to study the past in entirely new ways.

## References

- About the Polar Bear Expedition Digital Collection. (2015). Polar Bear Expedition Digital Collections. *Bentley Historical Library*. Retrieved from <http://quod.lib.umich.edu/p/polaread/about.html>
- Allison-Bunnell, J. (2009). Northwest digital archives: Evolving access to archives and special collections in the northwest. *OLA Quarterly*, 15(1), 31-35. Retrieved from [https://www.olaweb.org/assets/documents/olaq\\_15no1.pdf#page=33](https://www.olaweb.org/assets/documents/olaq_15no1.pdf#page=33)
- Allison-Bunnell, J., Yakel, E. & Hauck, J. (2011). Researchers at work: Assessing needs for content and presentation of archival materials. *Journal of Archival Organization*, 9, 67-104. Doi: 10.1080/15332748.2011.598400
- Amazon EFS preview. (2015). *Amazon Web Services*. Retrieved from <http://aws.amazon.com/efs/>
- Amazon S3. (2015). *Amazon Web Services*. Retrieved from <https://aws.amazon.com/s3/>
- Anderson, I. G. (2004). Are you being served? Historians and the search for primary sources. *Archivaria*, 35(3/4), 58-129. Retrieved from <http://journals.sfu.ca/archivar/index.php/archivaria/article/download/12479/135>
- Anderson, D., Kisiel, J., & Storksdieck, M. (2006). Understanding teachers' perspectives on field trips: Discovering common ground in three countries. *Curator*, 4, 365-386. Doi: 10.1111/j.2151-6952.2006.tb00229.x
- Archer, J., Hanlon, A., & Levine J. A. (2009). Investigating primary source literacy. *Journal of Academic Librarianship*, 35, 410-420. Doi:10.1016/j.acalib.2009.06.017
- Arthur, K., Byrne, S., Long, E., Montori, C. Q., & Nadler, J. (2004). Recognizing digitization as a preservation reformatting method. *Association of Research Libraries*. Retrieved from [https://siarchives.si.edu/sites/default/files/pdfs/digi\\_preserv.pdf](https://siarchives.si.edu/sites/default/files/pdfs/digi_preserv.pdf)
- Bantin, J., & Agne, L. (2010). Digitizing for value: A user-based strategy for university archives. *Journal of Archival Organization*, 8, 244-250. Doi:10.1080/15332748.2010.550791
- Barth, C. D. (1997). Archivists, genealogists, access and automation: Past and present trends in archival access technologies and their implications for the future of genealogical research in archives. *Infoneer.net*. Retrieved from <http://infoneer.net/cdb/archives1.shtml>
- Benedetti, J. M. (2003). A survey of small art museum libraries. *Art Documentation: Journal of the Art Libraries Society of North America*, 22, 31-39. Retrieved from <http://www.jstor.org/stable/27949263>

- Bik, H. M. & Goldstein, M. C. (2013). An introduction to social media for scientists. *PLoS Biology*, *11*(4), 1-8. Doi: 10.1371/journal.pbio.1001535
- Breakell, S. (2010). Evolving archival interfaces and the University of Brighton Design Archives. *Art Libraries Journal*, *35*(4), 12-17. Retrieved from <http://eprints.brighton.ac.uk/10018/>
- Burckel, N. C. & Cook, J. F. (1982). A profile of college and university archives in the United States. *The American Archivist*, *45*, 410-428. Retrieved from <http://www.jstor.org/stable/40292532>
- Collection use and extent statistics. (2013). *California Digital Library*. Retrieved from [http://www.cdlib.org/services/access\\_publishing/dsc/contribute/use.html](http://www.cdlib.org/services/access_publishing/dsc/contribute/use.html)
- Conway, P. (2010). Preservation in the age of Google: Digitization, digital preservation and dilemmas. *Library Quarterly: Information, Community, Policy*, *80*, 61-79. Doi: 10.1086/648463
- Chute, T. G. (2002). What's in a name? Outreach vs. basic services: A survey of college and university archivists. *Journal of Archival Organization*, *1*, 5-40. Doi: 10.1300/J201v01n02\_02
- Commonly asked questions about the use of school buses on field trips. (2012). The Highlands Districts Schools Transportation Department. Retrieved from <https://www.highlands.k12.fl.us/SBHC/Transportation/busFieldtripFAQ.pdf>
- Craig, B. L. (2000). A look at the bigger picture: The demographic profile of archivists in Canada based on a national survey. *Archivaria*, *1*(49), 20-52. Retrieved from <http://journalssfu.helloefei.com/archivar/index.php/archivaria/article/download/12738/13919>
- Cruikshank, K., Daniels, C., Meissner, D., Nelson, N. L., & Shelstad, M. (2005). How do we show you what we've got?: Access to archival collections in the digital age. *Journal for the Association for History and Computing*, *8*(2). Retrieved from <http://quod.lib.umich.edu/j/jahc/3310410.0008.203/--how-do-we-show-you-what-weve-got-access-to-archival?rgn=main;view=fulltext>
- Daines, J. G. III, & Nimer, C. L. (2009). *The interactive archivist: Case studies in utilizing web 2.0 to improve the archival experience*. Retrieved from <http://interactivearchivist.archivists.org>
- Dembo, J., & Custer, M. (2010). An experiment to increase online archival accessibility: using unique page views to measure online efficiency. *North Carolina Libraries*, *68*(2), 2-11. Retrieved from <http://www.ncl.ecu.edu/index.php/NCL/article/view/325>
- Dooley, J. M., Beckett, R., Cullingford, A., Sambrook, K., Sheppard, C., & Worrall, S. (2013). *Survey of special collections in the United Kingdom and Ireland*:

- Summary and recommendations*. Dublin, Ohio: OCLC Research. Retrieved from <http://www.oclc.org/content/dam/research/publications/library/2013/2013-01-sumrecs.pdf>
- Dooley, J. M., & Luce, K. (2010). *Taking our pulse: The OCLC research survey of special collections and archives*. Dublin, Ohio: OCLC Research. Retrieved from <http://oclc.org/content/dam/research/publications/library/2010/2010-11.pdf>
- Dowler, L. (1988). The role of use in defining archival practice and principles: A research agenda for the availability and use of records. *The American Archivist*, 51(1/2), 74-86. Retrieved from <http://www.jstor.org/stable/40293198>
- Dryden, J., Limkilde, C., Cherry, J. & Bogomazova, E. (2008). Archivists' views of user-based evaluation: Benefits, barriers and requirements. *The American Archivist*, 71, 144-166. Doi: 10.17723/aarc.71.1.y70837374478t146
- Duff, W., Dryden, J., Limkilde, C., Cherry, J., & Bogomazova, E. (2008). Archivists' views of user-based evaluation: Benefits, barriers, and requirements. *The American Archivist*, 71, 144-166. Doi: 10.17723/aarc.71.1.y70837374478t146.
- Duff, W. M., & Johnson, C. A. (2002). Accidentally found on purpose: Information-seeking behavior of historians in archives. *Library Quarterly*, 72, 472-496. Retrieved from <http://www.jstor.org/stable/40039793>
- Duff, W. M., & Johnson, C. A. (2013). Where is the list with all the names? Information-seeking behavior of genealogists. *The American Archivist*, 66, 79-95. Doi: /10.17723/aarc.66.1.l375uj047224737n
- Erway, R., & Schaffner, J. (2007). *Shifting gears: Gearing up to get into the flow*. Dublin, Ohio: OCLC Research. Retrieved from [www.oclc.org/programs/publications/reports/2007-02.pdf](http://www.oclc.org/programs/publications/reports/2007-02.pdf)
- Figuring the cost of a field trip. (2009). *Akron Public Schools* Retrieved from <http://old.akronschoools.com/departments/business-affairs/transportation-services/cost-of-field-trip.dot>
- Freeman, E. T. (1984). In the eye of the beholder: Archives administration from the user's point of view. *The American Archivist*, 47, 111-123. Retrieved from <http://www.jstor.org/stable/40292652>
- General Press Kit. (2015). About the Huntington. *The Huntington*. Retrieved from <http://www.huntington.org/WebAssets/Templates/content.aspx?id=11234>
- Gerken, J., Heilig, M., Jetter, H.C., Rexhausen, S., Demarmels, M., Konig, W. A., & Reiterer, H. (2009). Lessons learned from the design and evaluation of visual information-seeking systems. *International Journal on Digital Libraries*, 10(2), 49-66. Doi: 10.1007/s00799-009-0052-6

- Gilliland-Swetland, A. J. (1998). An exploration of K-12 user needs for digital primary source materials. *The American Archivist*, 61, 136-157. Retrieved from <http://www.jstor.org/stable/40294079>
- Goggin, J. (1986). The indirect approach: A study of scholarly users of black and women's organizational records in the Library of Congress Manuscript Division. *The Midwestern Archivist*, 11, 57-67. Retrieved from <http://www.jstor.org/stable/41057922>
- Gracy, K. F., & Kahn, M. B. (2011). Preservation in the digital age: A review of preservation literature, 2009-10. *Library Resources & Technical Services*, 56, 25-42. Retrieved from <https://journals.ala.org/lrts/article/view/5495/6745>
- Green, M. A., & Meissner, D. (2005). More product, less process: Revamping traditional archival processing. *The American Archivist*, 4, 208-263. Doi: 10.17723/aarc.68.2.c741823776k65863
- Greene, J. P., Kisida, B. & Bowen, D. H. (2014). The educational value of field trips: Taking students to an art museum improves critical thinking skills, and more. *Education Next*, 14(1), 79-86. Retrieved from <https://www.questia.com/library/journal/1G1-352376582/the-educational-value-of-field-trips-taking-students>
- Gueguen, G. (2010). Digitized special collections and multiple user groups. *Journal of Archival Organization*, 8, 96-109. Doi:10.1080/15332748.2010.513324
- Harvey, R. (2010). *Digital curation*. New York: Neal-Schuman Publishers.
- Hatton, G. (2006). Creating and managing archives for local history groups. *Historical Methods*, 39, 3-9. Doi:10.3200/HMTS.39.1.3-9
- Hecht, T. (2014). *Dancing archives – archive dances: Exploring dance histories at the Radcliffe College Archives*. Bielefeld, Germany: Transcript-Verlag.
- Herubel, J. P. V. M., & Goedecken, E. A. (1993). Trends in historical scholarship as evidenced in *The American Historical Review* 1896-1990. *Serials Review*, 19, 79. Doi: 10.1080/00987913.1993.10764152
- Hughes, L. M. (2003). *Digitizing collections: Strategic issues for the informed manager*. London: Facet Publishing.
- Hunter, G. S. (2003). *Developing and maintaining practical archives*. New York: Neal-Schuman Publishers.
- Huwe, T. K. (2009). Exploiting synergies among digital repositories, special collections and online community. *Online*, 33(2) 14-19. Retrieved from <http://cat.inist.fr/?aModele=afficheN&cpsidt=21178652>

- Increasing online access. (n. d.). *San Diego Air and Space Museum*. Retrieved from <http://sandiegoairandspace.org/research/project/increasing-online-access>
- Institute of Museum and Library Services. (2006). *Status of technology and digitization in the nation's museums and libraries*. Retrieved from [http://www.ims.gov/assets/1/AssetManager/Technology\\_Digitization.pdf](http://www.ims.gov/assets/1/AssetManager/Technology_Digitization.pdf)
- Institute of Museum and Library Services. (2014a). *Distribution of museums by discipline, FY 2014*. Retrieved from [http://www.ims.gov/assets/1/AssetManager/MUDF\\_TypeDist\\_2014q3.pdf](http://www.ims.gov/assets/1/AssetManager/MUDF_TypeDist_2014q3.pdf)
- Institute of Museum and Library Services. (2014b). *Museum universe data file*. Retrieved from [http://www.ims.gov/research/museum\\_universe\\_data\\_file.aspx](http://www.ims.gov/research/museum_universe_data_file.aspx)
- Kolowich, S. (2011, August 22). What students don't know. *Inside Higher Ed*. Retrieved from [https://www.insidehighered.com/news/2011/08/22/erial\\_study\\_of\\_student\\_research\\_habits\\_at\\_illinois\\_university\\_libraries\\_reveals\\_alarmingly\\_poor\\_information\\_literacy\\_and\\_skills](https://www.insidehighered.com/news/2011/08/22/erial_study_of_student_research_habits_at_illinois_university_libraries_reveals_alarmingly_poor_information_literacy_and_skills)
- Krause, M. G., & Yakel, E. (2007). Interaction in virtual archives: The Polar Bear Expedition digital collections next generation finding aid. *The American Archivist*, 70, 282-314. Doi: 10.17723/aarc.70.2.lpq61247881t10kv
- Kurutz, G. F. (2007). *Gold Rush*. California State Library. Retrieved from [www.library.ca.gov/goldrush/index.html](http://www.library.ca.gov/goldrush/index.html)
- Leboff, G. (2011). *Sticky marketing: Why everything in marketing has changed and what to do about it*. London: Kogan Page Limited.
- Lewis, B., & Griffin, M. (2011). Special collections and the new web: Using Libguides to provide meaningful access. *Journal of Electronic Resources Librarianship*, 23, 20-29. Doi: 10.1080/1941126X.2011.551091
- Lockers, M. L., Sag, M., & Schultz, J. (2012). Digital archives: Don't let copyright block data mining. *Nature*, 490, 29-30. Doi: 10.1038/490029a
- Los Angeles Unified School District. (2014). Charges for field trips. Retrieved from <http://achieve.lausd.net/Page/1796>
- Malkmus, D. K. (2008). Teaching undergraduates with primary sources. *Society of American Archivists -- 2008 Research Forum*. Retrieved from [http://www.crlt.umich.edu/sites/default/files/resource\\_files/Malkmus%20Teaching%20Undergraduates%20with%20Primary%20Sources.pdf](http://www.crlt.umich.edu/sites/default/files/resource_files/Malkmus%20Teaching%20Undergraduates%20with%20Primary%20Sources.pdf)
- Marchionni, P. (2009). Why are users so useful? User engagement and the experience of the JISC digitisation programme. *Ariadne: A Web & Print Magazine of Internet*



- Issues for Librarians & Information Specialists*, 61. Retrieved from [www.ariadne.ac.uk/issue61/marchionni](http://www.ariadne.ac.uk/issue61/marchionni)
- McCausland, S. (2011). A future without mediation? Online access, archivists, and the future of archival research. *Australian Academic & Research Libraries*, 42, 309-319. Doi: 10.1080/00048623.2011.10722243
- McFarland, C. (2007). Rethinking the business of small archives. *Archival Issues*, 31, 137-149. Retrieved from <http://www.jstor.org/stable/41102155>
- McKay, A. C. (2002). Genealogists and records: Preservation, advocacy, and politics. *Archival Issues*, 27, 23-33. Retrieved from <http://www.jstor.org/stable/41102053>
- Meissner, D., & Greene, M. A. (2010). More application while less appreciation: The adopters and antagonists of MPLP. *Journal of Archival Organization*, 8, 174-226. Doi:10.1080/15332748.2010.554069
- Monks-Leeson, E. (2011). Archives on the internet: Representing contexts and provenance from repository to website. *The American Archivist*, 74, 38-57. Doi: 10.17723/aarc.74.1.h386n333653kr83u
- Newell, J. (2012). Old objects, new media: Historical collections, digitization and affect. *Journal of Material Culture*, 17, 287-306. Doi: 10.1177/1359183512453534
- Northwest Digital Archives Usability Testing Working Group. (2006). *Usability principles*. Retrieved from <http://orbiscascade.org/index/cms-filesystem-action?file=nwda/reports/usabilityprinciplesutwg.pdf>
- Orion, N., & Hofstein, A. (1994). Factors that influence learning during a scientific trip in a natural environment. *Journal of Research in Science Teaching*, 29, 1097-1119. Doi: 10.1002/tea.3660311005
- Pampel, H. & Dallmeier-Tiessen, S. (2014). Open research data: From vision to practice. In S. Bartling and S. Friesike (Eds.), *Opening Science – The evolving guide on how the web is changing research, collaboration and scholarly publishing*. Retrieved from [http://book.openingscience.org/vision/open\\_research\\_data.html](http://book.openingscience.org/vision/open_research_data.html)
- Peacock, D., & Brownbill, J. (2007). Audiences, visitors, users: Reconceptualising users of museum on-line content and services. In J. Trant and D. Bearman (Eds.), *Museums and the Web 2007: Proceedings*. Toronto: Archives & Museum Informatics. Retrieved from <http://www.archimuse.com/mw2007/papers/peacock/peacock.html>
- Plumas County, California. (n.d.). Retrieved August 31 2015 from Wikipedia: [https://en.wikipedia.org/wiki/Plumas\\_County,\\_California](https://en.wikipedia.org/wiki/Plumas_County,_California)
- Plumas County Museum. (n.d.). About us. Retrieved from <https://www.plumasmuseum.org/about.html>

- Prom, C. J. (2012). Optimum access? Processing in college and university archives. *The American Archivist*, 73, 146-174. Doi:10.17723/aarc.73.1.519m6003k7110760
- Puglia, S., Reed, J., & Rhodes, E. (2004). *Technical guidelines for digitizing archival materials for electronic access*. Retrieved from <http://www.archives.gov/preservation/technical/guidelines.html>
- Puhek, M., Perse, M., & Sorgo, A. (2012). Comparison between a real field trip and a virtual field trip in a nature preserve: Knowledge gained in biology and ecology. *Journal of Baltic Science Education*, 11, 164-174. Retrieved from *EBSCOhost*
- Purcell, K., Rainie, L., Heaps, A., Buchanan, J., Friedrich, L., Jacklin, A. Chen, C., & Zichuhr, K. (2012). How teens do research in the digital world. Retrieved from <http://files.eric.ed.gov/fulltext/ED537513.pdf>
- Rose-Wiles, L., & Hofmann, M. A. (2013). Still desperately seeking citations: Undergraduate research in the age of web-scale discovery. *Journal of Library Administration*, 53, 147-166. Doi:10.1080/01930826.2013.853493
- Roth-Katz, E. (2012). Access and availability: A study of use policies on art museum library websites. *Art Documentation*, 31, 123-140. Doi: 10.1086/665335
- Scott County School District. (2014). Field trip request procedures. Retrieved from <http://www.scott.kyschools.us/docs/FY15%20FIELD%20TRIP%20TRANS%20COST%20INFO.pdf>
- Showley, R. (2007, August 12). Smokestacks & geraniums authors attempt to fill gaps in the local record. *Union Tribune San Diego*. Retrieved from [http://www.utsandiego.com/uniontrib/20070812/news\\_lz1h12histori.html](http://www.utsandiego.com/uniontrib/20070812/news_lz1h12histori.html)
- Smith-Yoshimura, K. & Cellentani, D. (2007). *RLG programs descriptive metadata practices survey results: Data supplement*. Dublin, Ohio: OCLC Research. Retrieved from <http://www.wip.oclc.org/content/dam/research/publications/library/2007/2007-04.pdf>
- Smith-Yoshimura, K., Godby, C. J., Koffler, H., Varnum, K., & Yakel, E. (2011). *Social metadata for libraries, archives and museums. Part 2: Survey analysis*. Dublin, Ohio: OCLC Research. Retrieved from <http://www.oclc.org/research/publications/library/2011/2011-03.pdf>
- Snow, K., Ballaux, B., Christensen-Dalsgaard, B., Hofman, H., Hansen, J. H., Innocenti, P., Neilsen, M. P., Ross, S., & Thogersen, J. (2008). Considering the user perspective: Research into usage and communication of digital information. *D-Lib Magazine*, 14(5/6). Retrieved from <http://www.dlib.org/dlib/may08/ross/05ross.html>

- Society of American Archivists. (2010). An introduction to the SAA. Retrieved from <http://www2.archivists.org/about/introduction-to-saa>
- Society of American Archivists. (2015). Glossary of archival and records terminology. Retrieved from <http://www2.archivists.org/glossary/terms/f/finding-aid>
- Southern Historical Collection. (n. d.). The Louis Round Wilson Library. <http://library.unc.edu/wilson/shc/>
- Stieg Dalton, M., & Charnigo, L. (2004). Historians and their information sources. *College and Research Libraries*, 65, 400-425. Doi: 10.5860/crl.65.5.400
- Sutton, S. C. (2012). Balancing boutique-level quality and large-scale production: The impact of "More Product, Less Process" on digitization in archives and special collections. *RBM: A Journal of Rare Books, Manuscripts, & Cultural Heritage*, 13, 50-63. Retrieved from <http://rbm.acrl.org/content/13/1/50.short>
- Tibbo, H. R. (2003). Primarily history in America: How U.S. Historians search for primary materials at the dawn of the digital age. *The American Archivist*, 66, 9-50. Doi: 10.17723/aarc.66.1.b1203701g718n74
- Van Ness, C. (2010). Much ado about paper clips: "More Product, Less Process" and the modern manuscript repository. *The American Archivist*, 73, 129-145. Retrieved from <http://www.jstor.org/stable/27802718>
- Velios, A. (2011). Creative archiving: A case study from the John Latham Archive. *Journal of the Society of Archivists*, 32, 255-271. Doi: 10.1080/00379816.2011.619705
- Walch, V. I. (2006a). Archival Census & Education Needs Survey in the United States (A\*CENSUS). *The American Archivist*, 69, 294-618. Retrieved from <http://www2.archivists.org/sites/all/files/ACENSUS-Final.pdf>
- Walch, V. I. (2006b). Part 3. A\*CENSUS: A closer look, expanded version. Retrieved from <http://www2.archivists.org/initiatives/acensus-archival-census-education-needs-survey-in-the-united-states>
- West, T., & Fesenko, K. (2009). Extending the reach of southern sources: Proceedings to large-scale digitization of manuscript collections. Final grant Report. Retrieved from [http://www2.lib.unc.edu/dc/watson/archivalmassdigitization/download/extending\\_the\\_reach.pdf](http://www2.lib.unc.edu/dc/watson/archivalmassdigitization/download/extending_the_reach.pdf)
- Wheeler, G. (2011). Field trips: Tradition in jeopardy. *Education Digest*, 76(8), 52-55. Retrieved from <http://eric.ed.gov/?id=EJ926287>
- Yakel, E. (2002). Listening to users. *Archival Issues*, 26, 111-127. Retrieved from <http://www.jstor.org/stable/41102044>

- Yakel, E., & Bost, L. L. (1994) Understanding administrative use and users in university archives. *The American Archivists*, 57, 496-615. Retrieved from <http://www.jstor.org/stable/40293870>
- Zastrow, J. (2013). The Opinion Piece: Digital Changes Everything: The Intersection of Libraries and Archives. *Computers in libraries*, 33(9), 16-18. Retrieved from <http://cat.inist.fr/?aModele=afficheN&cpsidt=27978678>
- Zhang, J. (2012). Archival representation in the digital age. *Journal of Archival Organization*, 10, 45-68. Doi: 10.1080/15332748.2012.677671

## **Appendix A: Haun Collection Test Site Assessment**

Page 1: Agreement to participate in research

Responsible Investigator(s): Lindsay Morton, Graduate Student, San José State University

Title of Study: Identity and User Preference in the Presentation and Content of Digital Archives: A Study of the Plumas County Museum's Haun Collection

1. You have been asked to participate in a research study investigating the presentation of primary source documents in an online environment.
2. You will be asked to interact with a website and complete a short survey about your experience.
3. No risks are anticipated as a result of participating in this study.
4. No direct benefits to the participants are expected.
6. Although the results of this study may be published, no information that could identify you will be included.
7. No compensation is available.
8. Questions about this research may be addressed to primary researcher Lindsay Morton at 530/228.4713 or by email at [linds.morton@gmail.com](mailto:linds.morton@gmail.com).

Complaints about the research may be presented to Sandra Hirsh, Director of the School of Library and Information Science at San José State University, 408/924.2490.

Questions about a research subjects' rights, or research-related injury may be presented to Pamela Stacks, Ph.D., Associate Vice President, Graduate Studies and Research, at 408/924.2427.

9. No service of any kind, to which you are otherwise entitled, will be lost or jeopardized if you choose not to participate in the study.
10. Your consent is being given voluntarily. You may refuse to participate in the entire study or in any part of the study.

You have the right to not answer questions you do not wish to answer. If you decide to participate in the study, you are free to withdraw at any time without any negative effect on your relations with San José State University or with the Plumas County Museum.

11. Please keep a copy of this form for your own records. If you would like to have a copy of this form sent to you via email for your records, please message [hauncollection@gmail.com](mailto:hauncollection@gmail.com). By agreeing to participate in the study, it is implied that you have read and understand the above information. Please do not write any identifying information on the survey/questionnaire.

Page 2: About your Experience

1. About the collection.

How would you rate the usefulness of the collection?

- Extremely useful
- Very useful
- Useful
- Somewhat useful
- Not useful

2. About your experience with the website.

How likely would you be to use a website like this again in the future?

- Extremely likely
- Likely
- Somewhat likely
- Somewhat unlikely
- Very unlikely

Page 3

3. Access to versions.

Which of the three views of the collection did you see? (Mark all that apply)

- Version 1
- Version 2
- Version 3

4. Version usefulness.

Which of the three views of the collection did you find most *useful* to you?

- Version 1
- Version 2
- Version 3

5. Rank each of the three styles in order of your personal preferences.

- Version 1 [options: 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>]
- Version 2 [options: 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>]
- Version 3 [options: 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>]

6. What particular features did you find most useful?

- Version 1: [fill in the blank]
- Version 2: [fill in the blank]
- Version 3: [fill in the blank]

7. What did you dislike?

- Version 1: [fill in the blank]
- Version 2: [fill in the blank]
- Version 3: [fill in the blank]

8. Any additional comments about your experience?

[free text area]

Page 4: About you.

9. Would you describe yourself as a:

- Scholar
- Student
- Teacher
- Informational/Avocational Researcher
- Other

10. Have you worked with historic documents before?

- Yes (in hardcopy)
- Yes (online)
- Yes (both online and in hardcopy)
- No

11. How many times per year do you consult primary source documents?

- Less than 1
- 1 – 2
- 3 – 6
- 6 or more

12. Please select highest level of education completed.

- Some High School
- High School Diploma
- Associate Degree
- Bachelors Degree
- Masters Degree
- PhD
- Other (please specify) [fill in the blank]

## Appendix B: Survey responses to written questions<sup>7</sup>

### 6. What particular features did you find most useful?

#### Version 1 Comments: (19 total responses)

- “The real and upfront history via pictures”
- “Helpful overview to the collection with clear links to material”
- “Chronological format”
- “Very clear resolution of the articles and documents.”
- “The organization”
- “Fast loading images, easy to zoom in”
- “The quality of the digital images and the ability to zoom in”
- “Chronological organization”
- “Digital page”
- “The actual letters”
- “Pictures were big enough to read”
- “Direct reading of the manuscript, full size”
- “Just the ability to access online”
- “Straightforward, beautiful scans”
- “Presents a good overview”
- “Simplicity of content organization, easy to find specific dates. Fewer links between section heads and content - could immediately access content upon entering this version.”
- “Viewing of original materials in context adds authenticity”
- “Nice images”
- “I liked the cast of characters and the abstract, as well as the explanations of the series”

#### Version 2 Comments: (18 total responses)

- “Transcript side by side with scan of document”
- “The history of the diaries and the ability to see how it has changed”
- “Chronological format”
- “Helpful supporting narration.”
- “The organization paired with the addition of details”
- “Editorial descriptive entry for each item. Facsimilies are slow to load and zoom.”
- “The biographical information was quite useful”
- “Digital page and the ability to "flip" pages.”
- “Layout”
- “Combination of big pictures and text”
- “Just the ability to access online”

---

<sup>7</sup> All comments reproduced as received, complete with errors in grammar and spelling.



- “Transcription”
- “I love the verbatim transcriptions combined with the clear images and the easy navigation from page to page. I can scan the text and then refer to the image when needed or pull up the large resolution image for a closer look.”
- “Same as above” (Previous answer: “Presents a good overview“
- “The transcriptions are very useful. The organizational scheme makes searching for content easier if you know who/when you are looking for. Navigating within digitized content is easier than in Version I.”
- “Transcription is very helpful as original materials are hard to read.”
- “Liked the ease in being able to go to next page”
- “I liked including the images.”

Version 3 Comments: (22 total responses)

- “Corrected documents, but links were still available to the original pages.”
- “Map”
- “The pictures were kept”
- “Easier to understand”
- “Helpful supporting narration and more appealing interface.”
- “The story like presentation made the information more interesting”
- “Transcript ok but too easy to miss emendations or errors. Facsimiles are hard to load and zoom.”
- “Transcriptions along with the biographical information was quite useful”
- “Easier to bounce to investigate specific questions & topics”
- “Don't care for typed versions.”
- “Version I and II look identical, with identical information and presentation. While Versions I and II are much more category focused, Version III is better for a casual glance at the family and their history, as opposed to the prior two versions being more report-oriented.”
- “Did not like”
- “Good organization”
- “All the useful accompaniments”
- “Just the ability to access online”
- “Transcription, additional information”
- “The transcripts are, of course, less work and easy read or skimmed through. This, plus the bells and whistles, would be great for students or the general public.”
- “Presents most detail”
- “I liked the synopses above the transcription, and the transcription was easy to read because of the edited formatting. As in Version II, navigating within digitized content is easier than in Version I. The extra stuff (“People” and “Essentials”) is very useful.”
- “Added comments are very helpful.”
- “Nice to have transcription”

7. What did you dislike?

Version 1 Comments: (11 total responses)

- “Difficult to navigate quickly”
- “Pop up windows”
- “Pop up as a picture, a lot of clicking to get too it”
- “Only visuals, no supporting information or translations.”
- “The lack of details”
- “Very little biographical or background information to contextualize the images”
- “Took too many steps to navigate to other topics of interests beyond specific account”
- “No page flipping”
- “Versoins I and II look extremely similiar, and are hard to tell apart.”
- “Confusing without text”
- “Slow due to need to download the larger images, and 19th century handwriting is often a chore to read.”
- “N/A”
- “Navigation within digitized content was time-consuming (having to toggle back a web page to open the next digitized page).”
- “Difficult to read handwriting, difficulty of reading handwriting and spelling a distraction”
- “Wish you could easily go to next image”
- “It was hard to understand what I was looking at with descriptors like "series"-- that's true for v2 as well”

Version 2 Comments: (16 total responses)

- “Difficult to navigate, harder to understand collection in its entirety”
- “The same issues as version 1” (Version 1 response: “Pop up as a picture, a lot of clicking to get too it”
- “Text somewhat undecipherable.”
- “Nothing”
- “No transcriptions”
- “None”
- “Can't see any downsides, frankly. I like the pull down access of the different documents by year.”
- “N/A”
- “No dislikes.”
- “All good”
- “Version 2 didn't have enough descriptive text--I wasn't sure what I was looking at/its significance”

Version 3 Comments: (16 total responses)

- “Difficult to navigate, harder to understand collection in its entirety”
- “The clicking various hyperlinks made the documents hader to find and when I did I had too many new tabs
- “Difficult to determine sequence of events
- “No complaints.

- “The lack of organization by category
- “See above” (Seems to refer to this text: “transcript ok but too easy to miss emendations or errors. Facsimiles are hard to load and zoom.”)
- “I would have loved to have the transcriptions directly next to the digitized images for comparison. Also I would have loved to know more about the account books such as what kinds of transactions they pertained to.”
- “Transcript.”
- “No next page button at bottom of page.”
- “Did not like the fact that they letters were edited.”
- “Too much text.”
- “Great as with ver 2, but as the transcription is not literal it is not entirely historically accurate although it is more accessible.”
- “N/A”
- “No dislikes.”
- “All good”
- “Not many complaints about version 3! Looks good.”

8. Any additional comments about your experience? (*14 total responses*)

- “Very neat project”
- “Well-formatted website, very intuitive and easy to access.”
- “After a quick survey I'm amazed at the work involved. This takes time and money. Most library and archival institutions I know are poorly funded and at best can afford only to develop finding aids. I would love to see every collection processed to this stage, but that's a fantasy I'm sure.”
- “I think Ms. Morton's digitization project is one that could be of superb quality and immensely useful with a few minor tweaks (as mentioned above).”
- “Although version three seems to be for the mass market, it may be too busy for a serious scholar only interested in the family's writings. Photos, reference maps, etc. helps peak interest of broader audience.”
- “Digital archives offer a lot of possible additions, but really, we're after the meat. Quick easy access to the digitized material itself. One should be able to flip through and know where you are.”
- “Awaiting Great War Era dairies and letters. Also, zoom on letters is extremely close. Perhaps a way to zoom in and out?”
- “none”
- “This is an amazing find. I have a historical family connection to the area. My great grandfather was a contemporary of David Haun and the Pratt family.”
- “I like the side by side comparisons of the three sites. I prefer versions 2 & 3”
- “I very much appreciate the abstracts of the texts which allow one to search for terms which are not explicitly stated in the documents. For example, out of curiosity, I search for "slave" and was able to find a letter describing the situation of several of the family's slaves even though the word "slave" never appeared in the letter itself. Also, the tags are useful as well. How about a family tree and/or glossary of names of people mentioned in the texts? Obviously, someone doing

full research with the letters is going to know that background, but students and other general users might benefit from a Who's Who, if you will.”

- “A very interesting study.”
- “Fabulous idea. Great job!”
- “Involving images on every page would probably make it more engaging--and replacing generic text with more descriptive titles would also help me understand what I was looking at.”