January 2008

Where's the Context? Enhancing Access to Digital Archives

Abigail R. Griner
University of Georgia

Where's the Context?
Enhancing Access to Digital Archives

Abigail R. Griner

I would contend that most objects of culture are . . . embedded within context and those contexts are embedded within other ones as well. So a characteristic of cultural objects is they're increasingly context-dependent.

-Brian Eno, *Time and Bits: Managing Digital Continuity*[1]

INTRODUCTION

Providing access to original materials is an ethical responsibility for all professional archivists. In the Code of Ethics for Archivists, access is the sixth tenet, stating that archivists not only provide equal and open access to records, they preserve the intellectual integrity of collections.[2] In an analog environment, this responsibility is somewhat straightforward

---

[1] Margaret MacLean and Ben H. Davis, *Time and Bits: Managing Digital Continuity* (Los Angeles: Getty Conservation Institute, Getty Trust), 1998, 51.

[2] "Code of Ethics for Archivists" (Chicago: Society of American Archivists, 2005), at <http://www.archivists.org/governance/handbook/app_ethics.asp> (accessed March 4, 2009).

and uncomplicated. However, technology has advanced rapidly over the past decade, and digitization projects are at the forefront of library and archival news. In a digital world, the once-simple tasks of promoting access to original materials and preserving their intellectual integrity are far more complicated. Although digitization has the potential to increase greatly a repository's patron base, complex decisions arise for archivists when contemplating this path. Institutions must expend more of their resources and staff to replicate digitally the value of analog collections. Many of these problems have been examined before, so I will address an issue that has been largely disregarded by archival literature: the necessity of placing digital collections within a broader social and historical context.

CONTRASTING DIGITAL AND ANALOG SETTINGS

Understanding context is vital for patrons researching archival collections. Unlike books, primary sources cannot stand by themselves. Thus, their level of description largely determines their long-term value. In the article "Archives Described at Collection Level," Meg Sweet and David Thomas state: "Archival documents can only be understood in the context in which they were created." Contextual information is also extremely critical when archival holdings contain sensitive subjects, topics that may be offensive to much of society now but were once acceptable. If understood in their proper historical context, these materials may not appear as offensive to researchers. Therefore, context is necessary for interpreting archival materials.[3]

Various kinds of contextual information may be obtained from archival collections. During their research, patrons learn about relationships between collections housed in the repository as well as in other institutions. They gather knowledge on historical trends, events, and figures related to the materials

---

[3] Meg Sweet and David Thomas, "Archives Described at Collection Level," *D-Lib Magazine* 6 (September 2000), <http://www.dlib.org/dlib/september00/sweet/09sweet.html> (accessed March 4, 2009); see also Abby Smith, *Why Digitize?*, (Washington, D.C.: Council on Library and Information Resources, 1999, 8-9), <http://www.clir.org/pubs/reports/pub80-smith/pub80.html> (accessed March 4, 2009); Michael Ester, *Digital Image Collections: Issues and Practices* (Washington, D.C.: Commission on Preservation and Access, 1996), 18.

they are studying. Before researchers even view an archival collection, a finding aid offers them descriptive information to place documents in context, which is vital for them to understand if a specific collection is relevant to their project. By adding historical context within finding aids, archivists already enhance access to analog collections. In addition to contextual information in finding aids, physically viewing original materials teaches researchers about the provenance of a collection and its connections to other people, places, and times.[4]

However, the research experience in a digital environment is entirely different from an analog setting. In an actual research room, users have the opportunity to examine whole boxes of materials, seeing the relationship between documents, folders, and series, and the correlation between these, the overarching collection, and even other collections held in the repository. The experience is very personal, and patrons often feel a strong connection to the physical materials. This does not occur in a digital environment, though. Researchers often find materials on the Internet by using a search engine, which leads them to the type of archival items they may or may not need without any way of showing how they reached them. Also, if users find digitized archives by browsing popular Web sites, they may not realize that certain images or documents have been decontextualized or misinterpreted. In most cases, even archival Web sites contain such minimal descriptive information that researchers could easily misinterpret their value or fail to see any relation to their studies. If digital archives do not provide patrons enough information to detail clearly the provenance and context of their holdings, the researchers will not be able to determine the reliability and quality of the evidence before them. In an analog setting, the researcher and archivist both have certain expectations and assumptions, but this is not true in a digital environment where archivists have no knowledge of who is viewing their collections, their level of research experience, or

---

[4] Aaron D. Purcell, "Providing Better Access to Manuscript Collections: A Case Study from the Historical Society of Washington, D.C.," *Journal of Archival Organization* 1 (2002): 37, 49.

the particular information for which they are looking.[5] Thus, the success of researchers in the digital setting depends even more on how well archivists do their jobs.

TRANSLATING ARCHIVAL COLLECTIONS ONLINE

Most archivists focus on the importance of provenance and chronology when creating their finding aids, providing detailed description of a collection as a whole, and only briefly summarizing individual series. This is a standard method and has been somewhat successful in an analog environment. However, it is not necessarily the best approach in a digital world. Even the *General International Standard for Archival Description,* or *ISAD(G)*, provides guidelines for archival description that do not always apply to digital spaces. Perhaps this is because *ISAD(G)* was developed at a time when digital space was first becoming a reality. For instance, the guidelines state that it is necessary to provide information relevant to the level of description. However, Abby Smith points out that online researchers want more information than most finding aids contain, especially in an environment where they cannot see the actual records and no reference archivist is readily available to assist them, as in a physical archives.[6]

Based on usability tests conducted at my institution, I have found that researchers tend to expect digital collections and finding aids to be more organized, better documented,

---

[5] Smith, *Why Digitize?*, 8-9; Bradley L. Schaffner, "Electronic Resources: A Wolf in Sheep's Clothing?," *College and Research Libraries* 62 (May 2001): 243; Christine L. Borgman, "The Invisible Library: Paradox of the Global Information Structure," *Library Trends* 51 (Spring 2003): 18-19, <http://www.ideals.uiuc.edu/bitstream/handle/2142/8487/librarytrendsv51i4j_opt.pdf?sequence=1> (accessed May 26, 2009); Sweet and Thomas, "Archives Described at Collection Level." Borgman, Schaffner, Sweet, and Thomas all discuss problems with search engines stripping context from digital items. Elizabeth Hallam Smith, "Lost in Cyberspace: Have Archives a Future?" (paper presented at the Australian Society of Archivists Conference, Melbourne, Australia, August 19, 2000), 10, <http://www.archivists.org.au/files/Conference_Papers/2000/hallamsmith.pdf> (accessed March 4, 2009).

[6] Sweet and Thomas, "Archives Described at Collection Level"; *ISAD(G): General International Standard Archival Description* (Ottawa: International Council on Archives, 2000), <http://www.ica.org/sites/default/files/isad_g_2e.pdf> (accessed May 26, 2009); Smith, *Why Digitize?*, 8-9.

and simpler to use than an actual physical archive. They want to understand fully the historical and social contexts of the collection materials they are browsing online. Avoiding redundancy of information is another *ISAD(G)* guideline that may need to be altered in a digital environment. Researchers may come across an archives page without understanding the path they took to get there. In order to avoid redundancy, an archivist might not have included contextual information on the accessed page or a link to it because the details are included on another Web page. But unless this is clearly stated, patrons may not understand it and fail to realize an item is relevant to their research.

It is important to note that much of the general public has extremely limited experience with archival sources, so few people have the research skills necessary to use primary sources effectively. But archival institutions still insist on digitizing collections for the Internet. Digitization allows researchers easier access to materials, but if they do not understand how to use original documents, digital archives will still not be an accessible research tool for them. Therefore, archival Web sites need to be simple for all user levels and include detailed explanations on their subject matter. Guidance on using archival collections or links to sites that provide tutorials on using archives would also add value. Otherwise, institutions are only reaching the same audience, those who already conduct research in a physical repository. In many cases, they are losing a younger, more computer-savvy group of potential patrons by failing to design user-friendly, archival Web sites.[7]

LITERATURE REVIEW

Although context is vital to understanding primary sources, many authors only briefly acknowledge the necessity of providing contextual information to digital collections. Diane Zorich's book, *Managing Digital Assets*, includes only two brief paragraphs on contextual information, referencing related technical issues. Donald Waters and John Garrett's 1996 volume does the same, but in more detail. Much literature focuses

---

[7] Abby Smith, *Strategies for Building Digitized Collections* (Washington, D.C.: Council on Library and Information Resources, 2001), <http://www.clir.org/pubs/reports/pub101/contents.html> (accessed March 4, 2009).

on contextual information in relation to metadata, though. Metadata is an excellent tool when digitizing collections, but researchers cannot see this information so it is not helpful for users who are trying to understand the social and historical context of materials. It is useful when implementing a searchable database of collections, but for patrons who would like to browse collections serendipitously, it is not a viable tool.[8]

Conversely, Stephen E. Ostrow acknowledges the importance of contextual information in relation to digital historical-image collections. He emphasizes the advantage of having a reading-room experience viewing photographs because researchers develop a greater understanding of a whole image collection by looking at folders within a box series, viewing groups of images at a time, and understanding their relation to each other and their role in the collection itself. Anne J. Gilliland-Swetland also discusses archival theory within a digital environment and the centrality of context but does not approach any specific problems associated with providing contextual information for digital archives. Still, she does successfully examine the disparity between concerns within the archival community and those in the library field in terms of digitization.[9]

Abby Smith gives the topic significant attention in two articles written for the Council on Library and Information Resources. Smith states that the analog and digital environment are significantly different, and a digital setting hinders researchers because a computer "flattens and decontextualizes" original

---

[8] Diane Zorich, *Managing Digital Assets: Options for Cultural and Educational Organizations* (Lose Angeles: Getty Information Institute, 1999), 62; Donald Waters and John Garrett, *Preserving Digital Information: A Report of the Task Force on Archiving Digital Information* (Washington, D.C.: Council on Library and Information Resources, 1996), <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf> (accessed May 26, 2009).

[9] Stephen E. Ostrow, *Digitizing Historical Pictorial Collections for the Internet* (Washington, D.C.: Council on Library and Information Resources, 1998), <http://www.clir.org/pubs/reports/ostrow/pub71.html> (accessed March 4, 2009); Anne J. Gilliland-Swetland, *Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment* (Washington, D.C.: Council on Library and Information Resources, 2000), <http://www.clir.org/pubs/reports/pub89/contents.html> (accessed March 4, 2009).

materials. She also maintains that archivists and librarians must carefully detail the digital collections they offer, even more so than analog materials. According to Smith, digitized items should actually be considered publications because they must be accompanied by an extensive amount of descriptive information in order to be understood in their broader historical and social context on the Internet. Aaron Purcell considers the issue as well, arguing that since digitization has become popular, archivists have focused on the technology issues associated with migrating archival materials to an electronic format, but in the process they have largely neglected content and context.[10]

TAKING A CUE FROM LIBRARIES

Perhaps the lack of archival literature on context and digital archives is related to the difference between perspectives in the library and archival fields. More libraries than archives have recently digitized their collections, particularly books and journals, but it does not necessarily hurt the value of these single-level items if context is not provided. Researchers may still gather quality information because they are meant to be examined as independent works. In contrast, archival collections are more valuable to patrons if viewed in terms of their provenance and historical context. Therefore, it seems the dire need for more literature on contextual information is related to the scarcity of resources for digitization projects in the archival world.

In many ways, archivists as well as librarians are still in the learning stages when it comes to digitization, and it is clear there are still no professional guidelines for certain areas of description for online collections. Libraries have more experience in digitization issues but library-and-information-science (LIS) theory is vastly different from archival theory. Although archives are generally studied in conjunction with LIS and history, archives in fact makeup a separate discipline with a unique body of theory, research, and professional experience. This can be detrimental to or work against expanding the archive research base. Thus, archivists need to develop their own digitization guidelines, and understand clearly the differences between digital libraries and digital archives. In considering

---

[10] Smith, *Strategies for Building Digitized Collections*; Smith, *Why Digitize?*, 8-9; Purcell, "Providing Better Access," 35.

this, archival institutions may begin to realize that the costs of digitizing archives are much higher in terms of time and resources than for creating digital libraries.

CHALLENGES TO PROVIDING CONTEXT

Several explanations may exist for the lack of contextual information on many archival Web sites. First, the nature of each collection is different, according to size, provenance, format, and research value. Most researchers would like all collections to be digitized, but this is not practical due to the lack of resources and funding within archival institutions. Therefore, archivists must select materials for digitization carefully. If an archivist selects a smaller, more manageable collection to digitize, it is generally easier to find contextual information because the description of each series is usually more detailed than that of a larger collection. This is not always the case, though, particularly if an archivist did not understand the research value of a collection at the time it was processed. In this case and that of other, larger archival collections, the lack of descriptive information will make it much more difficult for an archivist to provide context in an online environment. Also, to represent the content of larger collections, groups of individual documents or photographs are usually digitized instead of the entire collections. In this case, it is critical to provide contextual information since researchers are unable to compare all the records within series.

Deciding the amount of contextual information to include in a digital collection is a very difficult choice, and archivists must approach this on a case-by-case basis. According to Gilliland-Swetland, "the key is to explain the physical aspects and intellectual structure of the collection that may not be apparent and to provide enough contextual information for the user to understand the historical circumstances and organizational processes of the object's creation."[11] Some collections need little contextual information because the materials presented are fairly straightforward, particularly if they are small in size and created by a familiar individual or organization. Every archival institution should have a policy regarding their digital-collection presence and the inclusion of relevant contextual information should be detailed in this policy.

---

[11] Gilliland-Swetland, *Enduring Paradigm*.

For collections lacking contextual information within their finding aids, archivists need to perform more research to decide on an appropriate amount of information to add for digital reproductions. It is also vital for research work to determine contextual information to be done prior to or at the time of digitization. If not, vital information will be lost.[12] There are different methods of providing descriptive information other than rewriting current finding aids, though. Presenting a timeline of events relevant to the collection may be helpful. Users can then relate and compare items to each other and the larger collection as a whole in reference to the events described. Events on the timeline may be linked to a database detailing these topics. Links to people, place names, and images mentioned could also contribute in determining context. Linking to other similar records may be an option as well. When considering the importance of context, archivists must realize that ultimately it may be more practical to digitize more than less in many cases because researchers often draw context by seeing the relationship between records in a collection. Thus, archivists might consider digitizing collections that are related to one another or focus on some of the same topics. Therefore, regardless of the finding aid, additional contextual information may be identical for a certain group of collections.[13]

Employing any of these methods is quite labor intensive but the context it provides is very beneficial. Before digitization, archivists must understand the need for extremely descriptive information that details the context of archival materials. Their understanding of this will alter decisions when selecting materials because collections with limited background information will require much more time, effort, and resources for the institution.

---

[12] Waters and Garrett, *Preserving Digital Information*, 26; Ester, *Digital Image Collections*, 19; Kenneth Thibodeau, "Building the Archives of the Future: Advances in Preserving Electronic Records at the National Archives and Records Administration," *D-Lib Magazine* 7 (February 2001), <http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html> (accessed March 4, 2009).

[13] Samuel Gustman et al., eds., "Supporting Access to Large Oral History Archives" (paper presented at the International Conference on Digital Libraries, Portland, Ore., June 14-18, 2002), 9 <http://portal.acm.org/citation.cfm?doid=544220.544224> (accessed March 4, 2009); Ostrow, *Digitizing Historical Pictorial Collections*.

This may explain why many archival Web sites do not provide the information needed for researchers to understand the historical and social contexts of archival documents, photographs, and other materials. Archivists often fail to see the disparity between a research experience in an analog environment and a digital one. But they must learn methods to add value to digital items in order to produce digital collections more similar to analog records; providing contextual information is a significant way to do just that. The digital environment is changing the nature of research. We have a professional obligation to enable new types of research facilitated by a digital environment.

CONCLUSION

In order to determine the needs of researchers in an online environment, actual surveys should be conducted on user behavior on archival Web sites. Sweet and Thomas state that, "In practice many archive users require clear, accurate and searchable descriptions of individual files (or their equivalents). They then move 'bottom upwards' to see the context in which the documents were created and used."[14] This may or may not be true, but where is the documented research for this conclusion? And, if it is true, what should be the major priorities for archivists before posting digital collections to the Internet?

Archivists simply need to decide where their priorities lie and which ethical responsibilities are more important to them: providing equal access to online users and patrons in a physical archive or preserving the intellectual integrity of archival materials by including information that clearly communicates their historical and social contexts? (These may or may not be mutually exclusive.) Archival repositories hold valuable materials that the general public may have no knowledge of but which have the potential to make a great contribution to society. Thus, archival institutions have the ethical responsibility to disseminate this information to the public for the greater good. Otherwise, they will negate the potential of digital archives and their efforts will be for naught.[15]

---

[14] Sweet and Thomas, "Archives Described at Collection Level."

[15] Smith, "Lost in Cyberspace," 10.

Digitization in archives is often a choice between "depth and breadth."[16] Due to limited staff, resources, and time, many archival institutions end up choosing to digitize smaller collections in their entirety or a sizeable amount of materials within a range of large collections and including some contextual information from their current finding aids rather than expending time and effort to assess the finding aids to see if more research needs to be conducted in order to provide better description. Thus, quantity of digitized collections, not quality of information, becomes the priority. This is often a response to outside pressures from users demanding better access. Nonetheless, it is important to understand what kind of access is most beneficial to users instead of folding to impatient researchers.[17]

Abigail R. Griner is Access and Electronic Records Archivist at the Richard B. Russell Library for Political Research and Studies at the University of Georgia in Athens, Georgia. Before moving to Athens, she interned as a processing assistant at the Georgia Historical Society in Savannah, Georgia. She holds an M.A. in Public History from Armstrong Atlantic State University and an M.L.I.S. from the University of South Carolina.

---

[16] Sweet and Thomas, "Archives Described at Collection Level."

[17] Ibid.