

December 2015

Human Selection and Digitized Archival Collections: an Exploratory Research Project About Choice of Archival Materials Digitized for Online Public Availability

Randy Nelson Smith
University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Smith, Randy Nelson, "Human Selection and Digitized Archival Collections: an Exploratory Research Project About Choice of Archival Materials Digitized for Online Public Availability" (2015). *Theses and Dissertations*. 1101.
<https://dc.uwm.edu/etd/1101>

This Thesis is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

HUMAN SELECTION AND DIGITIZED ARCHIVAL COLLECTIONS: AN
EXPLORATORY RESEARCH PROJECT ABOUT CHOICE OF ARCHIVAL
MATERIALS DIGITIZED FOR ONLINE PUBLIC AVAILABILITY

by

Randy N. Smith

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Library and Information Science

at

The University of Wisconsin-Milwaukee

December 2015

ABSTRACT

HUMAN SELECTION AND DIGITIZED ARCHIVAL COLLECTIONS: AN EXPLORATORY RESEARCH PROJECT ABOUT CHOICE OF ARCHIVAL MATERIALS DIGITIZED FOR ONLINE PUBLIC AVAILABILITY

by

Randy N. Smith

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professor Donald Force

Our collective memory, the history that is cultivated through reflection, documentation, and consensus of historical data, is predicated upon the citizenry having access to the historical materials that society has created. Digitization has enabled greater public access to those materials. However, are items being scanned or digitally photographed to create surrogates that are then not made available to the world? The impetus for this study is to delve into whether or not intentional or unintentional personal choices play a role in determining which items archivists transform into digital surrogates; both in the decision of what to digitize and what to make available to the public on the World Wide Web. When one archival collection is prioritized over another or when it is not possible to digitize an entire collection, what rationale is used to determine which items will be digitized and published online? Do intentional or unintentional personal choices come into play in the decision-making? To answer these questions, four case studies were conducted, involving the random sampling of online collections and concomitant interviews of archivists. The purpose of this study is to enhance archivists' understanding of the reasons that guide the digitization decision-making process. Through such understanding, archivists can be more proactive in the decision-making process

to realize the benefit of digitizing and publishing archival materials that ultimately affect collective memory. The findings of this research revealed that in the case of the four institutions assessed, archivists do use personal choice to determine which materials within an archive are digitized.

© Copyright by Randy N. Smith, 2015
All Rights Reserved

To my wife, Leigh,
and kids, Calliope and Sagan,
for their patience, support and understanding;
my parents,
without whom this research would not have happened;
and Dr. Donald Force for all his
patience, guidance, and constant pushing toward excellence.

TABLE OF CONTENTS

	Abstract	ii
	List of Tables	viii
1	Introduction	1
	1.1 The digitization of archival materials	2
	1.2 Collective memory	6
	1.3 Digitization and collective memory	9
	1.4 Research questions	14
2	Literature Review	15
3	Methodology	24
4	Case Studies	29
	4.1 Institution A	29
	4.2 Institution B	33
	4.3 Institution C	35
	4.4 Institution D	37
5	Analysis	39
	5.1 Overview	39
	5.2 Collaboration	40
	5.3 Patron requests	42
	5.4 Student projects	43
	5.5 Exposure	44
	5.6 Administrative requests	44
	5.7 Preservation	45
	5.8 Donations	45
	5.9 Financial	45
6	Discussion	48
	6.1 Policies	48
	6.2 What's not digitized	50
	6.3 Evidence of personal choice	53
	6.4 Impact of the findings	54
	6.5 Study limitations	55
	6.6 Future research	57

7	Conclusion	59
	Bibliography	61
	Appendix A: Case Study Inquiry Email	68
	Appendix B: Interview Questions	70

LIST OF TABLES

Table 1. Comparison of the reasons for digitizing per institution	40
---	----

Chapter 1: Introduction

Archival materials that have been digitized, also known as *digital surrogates*,¹ provide greater public access² to primary and secondary source materials and can therefore help any one of various communities to define its collective memory. Although materials kept within an archival repository may be open to the public for research, that very accessibility might still be tantalizingly out of reach for the average person because of the time and money involved in physically visiting an archives,³ but it is possible to mitigate some of this lack of accessibility by way of having digital surrogates available on the World Wide Web.⁴ The impetus for this thesis is to delve into what choices archivists face — both in the decision of what to digitize and what to make available to the public on the World Wide Web — that play a role in determining which materials are transformed into digital surrogates and can thus affect collective memory.

¹ The phrase “digital surrogate” that will be used throughout this paper is understood to represent a digitized version of a physical item (such as a letter, photograph, receipt, audio recording, book, etc.) that has become a stand-in for the original item to both preserve the original item and potentially allow for wider public access. When referring to a digital surrogate I mean to reference each individual digital file created and stored on a computer medium, such as a hard disc, CD-ROM, flash drive, server, etc. The reason for this is to differentiate the scope of work involved; for example, the digitizing of ten single-page broadsides would yield a total of twenty digital files (a single file for each side of the broadside), whereas digitizing twenty different yearbooks with 200 pages each would yield a total of 4,000 digital image files. The term *digital surrogate* can also be found at the Archives Hub website (<http://archiveshub.ac.uk/glossary/#delta>).

² Laura Millar, *Archives : Principles and Practices*. (New York: Neal-Schuman Publishers, 2010), 199.

³ Maggie Dickson, “Due Diligence, Futile Effort: Copyright and the Digitization of the Thomas E. Watson Papers,” *American Archivist* 73, no. 2 (September 1, 2010): 626.

⁴ Access to the World Wide Web is still not universal. With regard to the United States alone, access to the Internet in 2012 was estimated to be about ninety-four percent according to the U. S. Federal Communications Commission. United States, Federal Communications Commission, *Eighth Broadband Progress Report*, (Washington D. C.), Federal Communications Commission, 2012: 23. Accessed November 7, 2015. https://apps.fcc.gov/edocs_public/attachmatch/FCC-12-90A1.pdf

When referring to “choices” or “personal choices” in this thesis, I am defining the phrase as the choices or values that *are not* based on professional policies, ethics or best practices. This definition of personal choice is contrary to my definition of professional choices which I define as the choices or values *are* based on professional policies, ethics, or best practices.

The outset of this chapter concerns where the archival field is in regard to mass digitization, providing insight for the reader into what practices are currently discussed as being utilized among archival institutions. This is followed by an exploration of what defines collective memory and some of the influences that affect it. The chapter concludes with a discussion about how digital surrogates could influence a community’s collective memory based on the decisions of those individuals digitizing archival materials.

1.1 The digitization of archival materials

Before materials are digitized, archivists must decide whether or not the digital surrogate will ultimately be accessible online. This decision is always relevant, but not always one that is considered by archivists. Digital surrogates may be created upon request and then stored locally or remotely on some form of closed media (i.e., optical discs, hard discs, etc.). Despite the benefit of access for digitizing archival materials, there are several reasons why an archives may not digitize archival materials, such as lack of time, donor restrictions, insufficient general funding, conservation concerns related to the materials, or not having the necessary technological infrastructure or financial means to utilize an outside digitization provider.⁵

⁵ Restrictions upon the publication of archival materials are a challenge that should be assessed on a case-by-case approach. Volunteers can often help with providing the human-power needed to perform a digitizing project.

The decisions about access to archival materials (specifically in the form of digital surrogates in this paper) have led to archivists being viewed as gatekeepers of the knowledge held within their repositories. Luke Gilliland-Swetland brings this to light when he discusses the historical development of the concept and principles of archival repositories, and the ‘archivist’ as a profession. Archivists have, by proxy, become the “gatekeepers” of the collections and materials under their purview.⁶ Additionally, Susan Pevar has described her own experience of becoming a “gatekeeper” to the collections held at the Langston Hughes Memorial Library at Lincoln University.⁷

Although archivists can be considered the experts of the collections under their purview, can they be considered experts on the historical contexts of which these materials are a part?⁸ The lack of historical expertise⁹ may impact the archivist’s choice about whether to digitize an item. For example, someone wanting to explore the travel habits of people involved in

Technological cost barriers are lowering. For example, the storage costs of digital files are lowering as resources open up to hold the digital surrogates with minimal or no cost to the repository. Online digital repositories are available for little or no cost such as the Internet Archive (<https://www.archive.org>), Flickr for digital images (<https://www.flickr.com>), or YouTube for digital video and audio recordings (<https://www.youtube.com>).

⁶ Luke Gilliland-Swetland, “The Provenance of a Profession: The Permanence of the Public Archives and Historical Manuscripts Traditions in American Archival History,” *American Archivist* 54, no. 2 (April 1, 1991): 160–75.

⁷ Susan Pevar, “Success as a Lone Arranger: Setting Priorities and Getting the Job Done,” *Journal of Archival Organization* 3, no. 1 (January 2005): 51–60, doi:10.1300/J201v03n105.

⁸ This definition of expert traces to Paul Conway’s article “Modes of Seeing: Digitized Photographic Archives and the Experienced User” Here, Conway draws upon the research presented in *How people learn: brain, mind, experience and school* describing an expert as having “acquired extensive knowledge that affects what they notice and how they organize, represent, and interpret information in their environment.” Paul Conway, “Modes of Seeing: Digitized Photographic Archives and the Experienced User,” *American Archivist* 73, no. 2 (Fall/Winter 2010): 425–62. (Bransford, Brown, and Cocking in Conway 2010, 432).

⁹ Peter Michel, “Digitizing Special Collections: To Boldly Go Where We’ve Been Before,” *Library Hi Tech* 23, no. 3 (2005): 381. Michel notes how a history graduate student was involved as part of a digitization project to provide the historical expertise needed to develop the context from the project.

nineteenth-century westward expansion might want online access to various ticket stubs, ledger books, receipts, etc. detailing the various costs, schedules, transportation systems, and obstacles involved. But these materials can be considered less visually appealing or interesting to a large portion of the population, and thus the impetus to digitize them may languish for decades with the items simply remaining untouched in the collection. An archivist may decide that an effort to digitize a photograph depicting a family sitting in pose is what would be of greater interest and appeal to the wider public. This personal assumption may end up being advantageous for the repository's status in that the digitized item might be well received by an online audience, or the digital surrogate could languish on a website somewhere, waiting to be accessed by someone browsing the Internet.

Access to digital surrogates also aids research through its broad — one might even say global — reach.¹⁰ The increased use of digital surrogates for public access to original source materials has been documented in the past,¹¹ demonstrating the reduction in research travel cost to access a remotely located item.¹² Such cost reductions can be a boon to research that utilizes primary source materials through remote access via the Internet.¹³ Because digitization

¹⁰ Ching-chih Chen et al., "Digital Imagery for Significant Cultural and Historical Materials," *International Journal on Digital Libraries* 5, no. 4 (August 1, 2005): 275–86, doi:10.1007/s00799-004-0097-5. Lisa Hooper and Donald C. Force, *Keeping Time : An Introduction to Archival Best Practices for Music Librarians* (Middleton, Wis: A-R Editions, Inc, 2014), 76, 78. Emily Monks-Leeson, "Archives on the Internet: Representing Contexts and Provenance from Repository to Website," *American Archivist* 74, no. 1 (Spring/Summer 2011): 41.

¹¹ Peter B. Hirtle, "The Impact of Digitization on Special Collections in Libraries," *Libraries & Culture* 37, no. 1 (Winter 2002): 43.

¹² Andrea Watson and P. Toby Graham, "CSS Alabama 'Digital Collection': A Special Collections Digitization Project," *The American Archivist* 61, no. 1 (April 1, 1998): 125.

¹³ Jasmine Elizabeth Burns, "Digital Facsimiles and the Modern Viewer: Medieval Manuscripts and Archival Practice in the Age of New Media" (M.A., State University of New York at Binghamton, 2013), 9, 23,

is about access, the greater the number of digitized materials that can be accessed, the greater the impact on research, in a clear quantitative approach.

Research utilizing digital surrogates, whether done by academics or by the general public, can be noted in several instances. Jasmine Burns,¹⁴ Jim Ginther,¹⁵ and I¹⁶ have all observed instances of the general public's aid in working with digital surrogates to further scholarship. However, if the archival material is not digitized, then access to it is restricted to physical visits.

Occasionally an archivist's desire to pursue the digitization of a particular part of a collection can result in the absence of items from the collection. For example, the George Engelmann Papers at the Missouri Botanical Garden¹⁷ are important to both botanists and historians of westward expansion during the nineteenth century. Engelmann was part of the nineteenth-century triumvirate of major botanists that included Asa Grey and John Torrey. The United States government and private botanists exploring the western states often coordinated

<http://search.proquest.com.ezproxy.lib.uwm.edu/pqdtft/docview/1431906037/abstract/CDF267A2DCA43A9PQ/1?accountid=15078>.

¹⁴ Ibid., 35. Burns discusses the work of Murtha Baca at the Getty Research Institute and her online tool to aid in the translation and analysis of seventeenth-century Spanish manuscripts.

¹⁵ Dr. Jim Ginther, at St. Louis University, has created an online tool called T-PEN (<http://t-pen.org/TPEN/>) to enlist the public's assistance with the transcription of manuscripts. Through personal communication, Dr. Ginther informed me that T-PEN has also been utilized by other institutions to aid in transcription.

¹⁶ Dr. Sven Eliasson, a retired professor and volunteer at the Peter H. Raven Library of the Missouri Botanical Garden (my employer), has been transcribing and translating many of George Engelmann's correspondences that have been uploaded to the Biodiversity Heritage Library (<http://www.biodiversitylibrary.org/collection/engelmannpapers>). Engelmann was originally from Germany, and while he also corresponded in English and French, a large number of the letters are in nineteenth-century German handwriting.

¹⁷ <http://www.biodiversitylibrary.org/collection/engelmannpapers>

with Engelmann on specimen-collecting. Engelmann also traded these specimen samples of western U.S. flora with botanical institutions and collectors from around the world. While a grant was pursued to digitize the correspondence part of the collection due to their historical significance, it was discovered toward the end of the project that Engelmann's scrapbook also contained letters; these were letters that Engelmann himself deemed important enough to set aside from the other correspondences and place in a scrapbook.¹⁸ How these letters within Engelmann's scrapbooks might alter the details of the U. S. westward botanical collection and exploration in general is unknown. However, the digitization and greater access to Engelmann's scrapbooks could alter the collective memory of this specific period in history.

Although digitization can aid in making primary source materials accessible to greater numbers of the online public, does this digitization have any effects on the culture of the public at large? For this part of the assessment we need to define the nature of the 'effect' on culture; however, to delineate that effect, an evaluation of collective memory must happen first. Once collective memory has been quantified it needs to be made clear how digitization interacts with collective memory.

1.2 *Collective memory*

To begin, a definition for 'collective memory' needs to be established, and for this, one can turn to University of Virginia history professor Alon Confino. In the article "Collective Memory and Cultural History: Problems of Method," he succinctly defines collective memory as

¹⁸ This example comes from first-hand experience, as I was a part of this project from its early stages through its conclusion. The scrapbook was never scanned as part of the project.

“the ways in which people construct a sense of the past.”¹⁹ However, Confino does further expand the definition by also stating, “It has been used to explore, first, the memory of people who actually experienced a given event, such as the memory of Holocaust survivors. In addition, it has come to denote the representation of the past and the making of it into a shared cultural knowledge by successive generations in ‘vehicles of memory’ such as books, films, museums, commemorations, and others.”²⁰

As Kerwin Klein notes, the connection between the public at large, both lay-people and professional academics, and the materials held within archival repositories has been a central issue since the early part of the twentieth century, when the interaction between collective memory and history first became part of scholarly discussions regarding society, repositories of history, and critical interpretation.²¹ Primary source materials can help a society retain collective memory; likewise, they help prevent collective memory from becoming whitewashed history²² beholden to political or ideological agendas.²³

¹⁹ Alon Confino, “Collective Memory and Cultural History: Problems of Method,” *The American Historical Review* 102, no. 5 (December 1, 1997): 1386, doi:10.2307/2171069.

²⁰ *Ibid.*, 1386.

²¹ Kerwin Lee Klein, “On the Emergence of ‘Memory’ in Historical Discourse,” *Representations*, no. 69 (January 2000): 127.

²² *Ibid.*, 130.

²³ Pierre Nora, “Between Memory and History: Les Lieux de Mémoire,” *Representations*, no. 26 (April 1, 1989): 7–24, doi:10.2307/2928520. Pierre Nora writes, “The indiscriminate production of archives is the acute effect of a new consciousness, the clearest expression of the terrorism of historicized memory.” His article, while directing the reader toward the beauty of memory versus the sterility of history, does admit that archives have been invented to maintain artifacts from the past to help our memories from straying too far from our history.

Collective memory has been explored, sometimes as an isolated case, as in the book *History Wars: The Enola Gay and Other Battles for the American Past*, which chronicles the backlash and politics involved with the Smithsonian's planned 1995 exhibit that presented a more critical look at the United States' dropping of atomic bombs on Hiroshima and Nagasaki,²⁴ and sometimes as an over-arching investigation, in many different manners. In dealing with the investigation, Confino analyzed the concept of collective memory and how it relates to cultural history believing collective memory to be made up of the relationships of the social, political, and cultural, and also the relationship between "representational and social experience."²⁵

Collective memory itself has been shown to be faulty²⁶ and sometimes in need of scholarship to step forward to challenge it. Or, as Pierre Nora writes, "memory, insofar as it is affective and magical, only accommodates those facts that suit it."²⁷ The theme of memory was further explored by Barbara L. Craig, whereby she advocates the need for archival collections to be accessible to help shape memory.²⁸ Gerald Sider demonstrated this in an article that recounted his investigation into the history of a 1912 textile strike in Lawrence, Massachusetts. He revealed how one journalist's interest in a story, his lack of rigorous research, and the

²⁴ Edward Tabor Linenthal and Tom Engelhardt, eds., *History Wars: The Enola Gay and Other Battles for the American Past*, 1st ed (New York: Metropolitan Books, 1996).

²⁵ Confino, "Collective Memory and Cultural History," 1402.

²⁶ Barbara L. Craig, "Selected Themes in the Literature on Memory and Their Pertinence to Archives," *American Archivist* 65, no. 2 (Fall 2002): 276–89.

²⁷ Nora, "Between Memory and History," 8.

²⁸ Craig, "Selected Themes in the Literature on Memory and Their Pertinence to Archives," 276-89.

promotion of a narrative based more on conjecture than fact led to an inaccurate historical revelation about a 1912 textile strike.²⁹

Along with its potential faults, collective memory also feeds off of itself through repetition of a narrative when coming from popular cultural sources. In an informal study, Michael Frisch notes how collective memory influences pedagogy, which in turn feeds back into the collective memory.³⁰ Along with Frisch's pedagogical influence, Emily Monks-Leeson has discussed the affects postmodernism is having on collective memory as the digital surrogates provide a new, more accessible way to develop contextual narratives.³¹

Collective memory has also been explored as different mechanisms of societal traditions. In the introduction to *The Invention of Tradition*, Eric Hobsbawm discusses the meaning of invented tradition in societies and its role in shaping collective memory. Hobsbawm notes that the invented traditions of nationalism "all rest on exercises in social engineering which are often deliberate and always innovative."³²

1.3 *Digitization and collective memory*

²⁹ Gerald M. Sider, "Cleansing History: Lawrence, Massachusetts, the Strike for Four Loaves of Bread and No Roses, and the Anthropology of Working-Class Consciousness," *Radical History Review* 1996, no. 65 (March 20, 1996): 48–83, doi:10.1215/01636545-1996-65-48.

³⁰ Michael Frisch, "American History and the Structures of Collective Memory: A Modest Exercise in Empirical Iconography," *The Journal of American History* 75, no. 4 (March 1, 1989): 1130–55, doi:10.2307/1908633.

³¹ Monks-Leeson, "Archives on the Internet."

³² E. J. Hobsbawm and T. O. Ranger, eds., *The Invention of Tradition* (Cambridge: Cambridge University Press, 1995), 13.

The placing of digital surrogates on the Internet presents a virtual public exhibit of the archival materials held at one or more repositories and has the potential to affect collective memory through virtual public exhibition.³³ Susan Crane and Ryan Davis have discussed the importance of this exhibited information (e.g., digital surrogates) impacting collective memory. Crane notes that antithetical information can have a shock effect on collective memory.³⁴ Davis further explored how collective memory is maintained by the impact of an online exhibit.³⁵

That is not to say that all archivists who select materials to be digitized and made public do so with an agenda in mind. Previous research has demonstrated the power of popular ideas (e.g., collective memory) that are reinforced despite the accuracy of the information.³⁶ Specific archival materials could possibly be viewed as more significant and in greater need of digitization based upon contemporary collective memory influence upon an archivist. For example, if a particular archival item or collection demonstrates a counter narrative to the collective memory, the archivist may choose not to digitize it to avoid any negative reactions.

This decision to avert negative reactions to digital surrogates by not selecting the original archival material is an example of controversy avoidance. Gary Nash, Charlotte

³³ Craig, "Selected Themes in the Literature on Memory and Their Pertinence to Archives."; Lilly Koltun, "The Promise and Threat of Digital Options in an Archival Age," *Archivaria*, no. 47 (April 15, 1999): 115. Koltun discusses how any context through exhibition can influence collective memory.

³⁴ Susan A. Crane, "Memory, Distortion, and History in the Museum," *History & Theory* 36, no. 4 (December 1997): 44.

³⁵ Ryan A. Davis, "Caught in Franco's Web: Virtual Memories of Prisons During the Franco Regime," *Journal of Spanish Cultural Studies* 9, no. 1 (2008): 90, doi:10.1080/14636200701868035.

³⁶ Sara M. Baker and Richard E. Petty, "Majority and Minority Influence: Source-Position Imbalance as a Determinant of Message Scrutiny," *Journal of Personality and Social Psychology* 67, no. 1 (July 1994): 5–19, doi:10.1037/0022-3514.67.1.5.

Crabtree, and Ross Dunn explored this issue by noting an example of textbook publishers. Textbook publishers have often faced the issue of introducing information that may offend and thus often had authors rewrite sections of the text, especially in history textbooks, to be appealing to as wide an audience as possible.³⁷ Textbooks, as with websites, serve to impart information to a general public, and by extension the same influences upon textbooks also influence which archival materials are available via the Internet.

Collective memory can benefit from digitization in two other ways: “new types of research” and new users and uses.³⁸ As Peter Hirtle notes, digitization has led to new techniques, such as utilizing computer enhancements to reveal artifacts that are beyond what is visible to the human eye, as well as being able to do such things as linking transcriptions and translations to a digital surrogate.³⁹ With the global reach of digital surrogates,⁴⁰ collective memory is also being shaped through new or renewed uses of the information provided from digital surrogates.⁴¹ A prime example of this is the United States Civil War website, *The Valley of the Shadow*,⁴² which through multi-disciplinary collaboration brought digitized materials

³⁷ Gary B. Nash, Charlotte A. Crabtree, and Ross E. Dunn, *History on Trial: Culture Wars and the Teaching of the Past*, 1st ed. (New York: Alfred A. Knopf, 1997): 70.

³⁸ Hirtle, “The Impact of Digitization on Special Collections in Libraries,” 44.

³⁹ *Ibid.*, 44.

⁴⁰ Leslie Johnston, “Making It Usable: Developing Personal Collection Tools for Digital Collections,” in *I, DIGITAL: Personal Collections in the Digital Era* (Society of American Archivists, 2011), 258-9.

⁴¹ Monks-Leeson, “Archives on the Internet,” 39, 54-5.

⁴² *The Valley of the Shadow: Two Communities in the American Civil War*: <http://valley.lib.virginia.edu>

together,⁴³ to create a website to help convey history to shape collective memory.⁴⁴ In this case, historians create context from the many digital surrogates⁴⁵ to shape collective memory. Another example is one that I have a personal connection with. Regarding the George Engelmann correspondence project mentioned earlier, Dr. Sven Eliasson,⁴⁶ who is transcribing and translating the letters, has personally communicated to me that he prefers using the digitized copies of Engelmann's letters over the originals as he can alter the digital images to help make the text in the letters easier to read. The translations and transcriptions that are completed by Dr. Eliasson are then linked online to the original letters. Regardless of the anecdotal example this illustrates, this situation demonstrates one user's feedback in utilizing a digital surrogate of an object rather than the original.

A research participant in a study by Paul Conway spoke about their selection of digitized photographs to illustrate a context they wished to promote. In reference to a photograph depicting a large outdoor gathering of people and horses near a building, the participant stated, "I'm selecting images that are conveying culturally particular aspects that the Russians seem to be honing in on [in assembling the album]."⁴⁷ And as Conway states, "for all of the participants,

⁴³ Edward L. Ayers, "Doing Scholarship on the Web: Ten Years of Triumphs - and a Disappointment," *Journal of Scholarly Publishing* 35, no. 3 (2004): 143–7.

⁴⁴ Koltun, "The Promise and Threat of Digital Options in an Archival Age," 115.

⁴⁵ Michel, "Digitizing Special Collections," 381. Michel describes that one of the goals of the digitization of the special collections at the University of Las Vegas was to demonstrate how the different collections were related.

⁴⁶ See note 17.

⁴⁷ Conway, "Modes of Seeing," 442.

if the image fit the purpose and the product, all that remained were technical decisions relating to publishing.”⁴⁸

In his article detailing three experiences he had with the community of Berisso, Argentina, Daniel James demonstrates how his experiences revealed the fluidity of collective memory. He describes how several forces, forces we might not necessarily view as very strong such as children or a news reporter, can begin to shape collective memory.⁴⁹ His argument reveals two separate facets about memory: firstly, even tiny pieces of information (e.g., digital surrogates) can exert force over collective memory, and secondly, that the lack of information can also exert force upon collective memory. An example of the second facet goes back to Conway’s study where he notes that research participant four adjusted his own research to accommodate the lack of a digital surrogate, stating, “I’m happy with what’s online and if it’s not online, I just try to work around it.”⁵⁰ As Lisa Hooper and Donald Force state, “researchers should be aware that there are likely many pertinent documents not included in [an] exhibition that could have some bearing on [a] narrative.”⁵¹

The public influence on collective memory can also come from the lack of expertise of the archivist. Such a case can simply be the identification, or re-identification, of data from an archival item. Conway notes how a research participant noted the misidentification of a

⁴⁸ Ibid., 442.

⁴⁹ Daniel James, “Meatpackers, Peronists, and Collective Memory: A View From the South,” *American Historical Review* 102, no. 5 (December 1997): 1404-12.

⁵⁰ Conway, “Modes of Seeing,” 454.

⁵¹ Hooper and Force, *Keeping Time*, 76.

baseball player within a digital surrogate of a photograph held at the Library of Congress⁵² and how another participant “found new connections” between previously separated photographs.⁵³ This point about the connection of digital surrogates is important because their placement online means that they are now part of the global network; the connections and context are created from digital surrogates.

1.4 *Research questions*

This thesis is being pursued to investigate how personal choices affect not only the selection of materials to be digitized but also decisions related to which digitized materials are made available online and their possible effect on collective memory. To that end, this research project is guided by the following questions: When one archival collection is prioritized over another or when it is not possible to digitize an entire collection, what rationale is used to determine which items will be digitized? Do intentional or unintentional personal selection biases come into play in the decision-making? Moreover, what are the reasons that archivists cite when choosing which digitized materials to present online? And do intentional or unintentional biases drive these choices?

⁵² Conway, “Modes of Seeing,” 439-40.

⁵³ *Ibid.*, 453.

Chapter 2: Literature Review

The topic of digitization of archival materials has been widely discussed by archivists. The discussions have ranged from the benefits of metadata for location and retrieval of digitized materials to the anecdotal case studies of digitization projects. Yet none that I have encountered have systematically and comprehensively studied the element of choices made by the archivist when deciding what should and should not be made publicly available via digital surrogates.

In a government-commissioned survey of research libraries and archives carried out to assess the state of digitization efforts in the United Kingdom, Barbara Bültmann, et al. addressed the selection of archival materials for digitization.⁵⁴ They found that selection of archival materials to be digitized fell into nine categories, with four of these categories standing out as the most prevalent: public demand for the material, institutional relevance, the uniqueness of the item(s), and whether or not a collection was “coherent.”⁵⁵ At the end of the article, the researchers draw some conclusions to aid in methodology for conducting a digitization project, but include only one instruction to aid in selecting archival materials for digitization: the identification of gaps within digitized collections and reference request.⁵⁶

⁵⁴ Barbara Bültmann, et al., “Digitized Content in the UK Research Library and Archives Sector,” *Journal of Librarianship and Information Science* 38, no. 2 (January 1, 2006): 106.

⁵⁵ *Ibid.*, 108-9. There is not a clear definition given for the term “coherent collection(s)” used in the survey. From this, I infer they mean that enough archival materials within the collection are available to present some form of narrative or relationship if digitized.

⁵⁶ *Ibid.*, 120.

Aside from the work of Barbara Bültmann et al.,⁵⁷ and Lisa Hooper and Donald Force (to be discussed in more detail later),⁵⁸ most of the literature presents the digitization selection process by turning it into a methodology. By this I mean that the methodological steps are linear in fashion so as to remove ambiguity from the process. For example, was a grant received to scan and publish online collection X? If yes, does the repository have all the applicable permissions in place to scan said collection? If yes, scan and publish online collection X, etc.

The National Archives and Records Administration (NARA) and the Canadian Council of Archives (CCA) have both published online works to aid in the selection of materials to be digitized.⁵⁹ But these resources, and others,⁶⁰ treat the selection and implementation of digitizing materials as a very structured process, which will be now summarized.

Digitization by request is often used as a prevalent means for determining what to digitize. Aside from NARA's and CCA's selection implementation plans, James Bantin and Leah Agne go one step further in exploring the issue of selection itself. They mention looking beyond prescriptive plans toward prioritizing reference requests to determine digitizing priority.⁶¹ It could be argued that quantifying the reference request to determine such priorities is itself an

⁵⁷ Ibid.

⁵⁸ Hooper and Force, *Keeping Time*.

⁵⁹ "National Archives and Records Administration," *Archives.gov*, May 20, 2014, <http://www.archives.gov/digitization/strategy.html>. The CCA even offers their selection process as a chart called a "Decision Tree for Digitization Projects" to assist with project development. Preservation committee, "Digitization and Archives" (Canadian Council of Archives, 2002), http://www.cdncouncilarchives.ca/digitization_en.pdf.

⁶⁰ It should be noted that NARA and the CCA are not the only entities to have discussed digitization selection plans. For example, Laura Millar has a few pages devoted to the process as well. Millar, *Archives*, 199-204.

⁶¹ James Bantin and Leah Agne, "Digitizing for Value: A User-Based Strategy for University Archives," *Journal of Archival Organization* 8, no. 3/4 (July 2010): 244-45, doi:10.1080/15332748.2010.550791.

added layer of structure upon the already existing ones being described here. As long as the repository has had a method to collect such data, this quantification can lend itself to an empirical layer of material selection.

If the repository does not have hard data on reference requests, sometimes a need arises for a committee (to enable input from multiple participants and thus reduce individual bias) to select what they consider to be most pertinent for public interest. Such a situation occurred at the University of Alabama when they received funding to digitize a collection, which led to debates about how to best aid and be “appealing” to school educators, while promoting awareness of the collection.⁶²

Akin to the concept of patron-driven choices, is patron-produced choices. Evan Robb’s article on community-based digitization discusses a project whereby the public brought their personal items deemed to be special in some personal respect to be digitized by Washington State Library.⁶³ In this manner, the choice of digitization is removed from the institution and placed completely on the public’s decision-making processes.

Another criterion present in the prescriptive plans are as follow: assurances have been obtained that the item is either out-of-copyright or has been given permission to be published. Melissa Levine⁶⁴ and Maggie Dickson⁶⁵ both look into the legal issues, including copyright,

⁶² Watson and Graham, “CSS Alabama ‘Digital Collection,’” 125.

⁶³ Evan Robb, “Gleaning Local History: Community-Based Digitization Experiences in Rural Washington,” *Microform & Imaging Review* 39, no. 1 (January 2010): 12–17, doi:10.1515/mfir.2010.002.

⁶⁴ Melissa Smith Levine provides a good synopsis of the legal issues involved with digitizing archival materials in her appropriately titled article “Overview of Legal Issues for Digitization.” Maxine K. Sitts, ed., *Handbook for Digital*

which need to be addressed before tackling a digitization project. This subject is further explored with an analysis of the copyright issues surrounding *Europeana*;⁶⁶ from the image, to the metadata, to the laws themselves.⁶⁷ The issue of copyright extends backwards from the digital surrogate to the selection of materials to be digitized.

Another criterion is that the equipment to be used is capable of digitizing the item (i.e., it would be difficult to digitize a slide with a digital SLR [single-lens reflex] camera). The personnel must be capable of implementing best-practice standards regarding both the digitizing itself and the metadata to be associated with the digital surrogate. And finally, the original source material needs to be hardy enough to withstand the potentially rigorous handling during the digitization procedure. These precautions not only apply to text-based materials, but also include moving images, audio recordings, and artifacts as well.⁶⁸

These are by no means the only suggestions for the selection of materials to be digitized. For example, Paul Conway's focused view of digitizing for preservation presents the somewhat vague selection criteria for digitizing those items that have the potential for

Projects : A Management Tool for Preservation and Access, 1st ed. (Andover, Mass: Northeast Document Conservation Center, 2000).

⁶⁵ Dickson, "Due Diligence, Futile Effort."

⁶⁶ <http://www.europeana.eu/portal/>

⁶⁷ Jonathan Purday, "Intellectual Property Issues and Europeana, Europe's Digital Library, Museum and Archive," *Legal Information Management* 10, no. 03 (2010): 174–80, doi:10.1017/S1472669610000678.

⁶⁸ Karen F. Gracy, "Distribution and Consumption Patterns of Archival Moving Images in Online Environments," *American Archivist* 75, no. 2 (Fall/Winter 2012): 446-7.

“increasing impact and distinctiveness at home.”⁶⁹ This idea is similar to that of special collections librarian Peter Michel, who, for a digitization project, suggested that the selection of materials to digitize was heavily influenced by the desire to simply “promote and highlight collections.”⁷⁰ Also, Ching-chih Chen et al. propose that digitization should focus on rare or priceless known works so that digital surrogates exist in the event the originals are destroyed.⁷¹

Lisa Hooper and Donald Force form four conditional criteria for the selection of materials to digitize: value, condition, access, and legal rights.⁷² Although their discussion of “value,” “condition,” and “legal rights” falls under the methodological categories discussed earlier, I believe it is the criterion of “access” that stands out as most unique. They specifically mention that digitizing “hidden” (quotes are theirs) materials helps illuminate those materials that otherwise may go unnoticed.⁷³ In an editorial for *Library Hi Tech*, Michael Seadle echoes Hooper and Force’s final idea of digitizing “low-value materials,” noting that physical time is the materials’ biggest threat from deterioration.⁷⁴ The deterioration of textual materials from the nineteenth-century is also the focus of an article by Andrew Stauffer.⁷⁵ He notes that without digitization to reduce physical use and provide wider access to the information, research

⁶⁹ Paul Conway, “Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas,” *Library Quarterly* 80, no. 1 (January 2010): 75.

⁷⁰ Michel, “Digitizing Special Collections,” 382.

⁷¹ Chen et al., “Digital Imagery for Significant Cultural and Historical Materials.”

⁷² Hooper and Force, *Keeping Time*, 78.

⁷³ *Ibid.*, 78. Monks-Leeson, “Archives on the Internet,” 41. Monks-Leeson also notes this same concept in her article.

⁷⁴ Michael Seadle, “Selection for Digital Preservation,” *Library Hi Tech* 22, no. 2 (2004): 120.

⁷⁵ Andrew Stauffer, “The Nineteenth-Century Archive in the Digital Age,” *European Romantic Review* 23, no. 3 (June 2012): 335–41.

involving these sources will be hampered. A very similar two-year project was undertaken by the Press Institute of Mongolia to primarily preserve Mongolian newspapers that were in the process of deteriorating but also allow greater access to underrepresented materials.⁷⁶ In this case, although the newspapers were printed between 1990 and 1995, many of the copies were not collected or well maintained⁷⁷; this situation demonstrates that not all archival materials in need of preservation and/or access are necessarily also antique in nature. The goals of the World Newspaper Archive seem to fall in line with the goals of the Press Institute. James Simon examines the World Newspaper Archive and states that the main goals are access, preservation, and growth (acquiring more digital surrogates of newspapers).⁷⁸ Burns states that, “institutions give precedent to the conservation of the object over access of the scholar,”⁷⁹ but beyond this, Burns is more interested in the pros and cons of digitization than the selection. This statement would seem to indicate that Burns feels that the preservation of materials often trumps the desires of patrons with regard to the selection of materials to be digitized.

Somewhat contrary to Burns’ hypothesis, Emily Monks-Leeson demonstrates how provenance and context can be maintained or transformed through a digitization project.⁸⁰

⁷⁶ Krystyna K. Matusiak and Myagmar Munkhmandakh, “A Newspaper/Periodical Digitization Project in Mongolia: Creating a Digital Archive of Rare Mongolian Publications,” *The Serials Librarian* 57, no. 1–2 (2009): 118–27, doi:10.1080/03615260802669136.

⁷⁷ *Ibid.*, 119.

⁷⁸ James T. Simon, “World Newspaper Archive,” *Slavic & East European Information Resources* 10, no. 1 (January 2009): 83–4, doi:10.1080/15228880902756443.

⁷⁹ Burns, “Digital Facsimiles and the Modern Viewer,” 41.

⁸⁰ Monks-Leeson, “Archives on the Internet,” 38-57.

Although she does discuss this very enlightening issue, she does not discuss how the selection of materials to create a digital surrogate can also affect context; the creation of a digital surrogate does not likely affect provenance as that issue is something that influences each surrogate individually. Similarly, by demonstrating the transformative scholarship achieved through the collaboration of three universities, Lorena Gauthereau-Bryson notes that context is better understood through a “hemispheric perspective” in which archival materials were digitized from the respective institutions in two countries.⁸¹

Bantin and Agne have noted that the amount of time that goes into the planning of digitization projects can be considerable.⁸² With limited financial or personnel resources, some repositories may find it difficult to be able to follow a methodical digitization plan and may simply need to resort to an ad hoc approach to digitization. To explore this financial issue in more detail, Pauline Joseph, Michael Gregg, and Sally May did a case study regarding a project to digitize materials relating to the anniversary of the Western Museum Welcome Wall in Australia and raise an excellent point: if this archival material has a pseudo-surrogate, that is a digital surrogate that is not a digitization of the original item (for example, a PDF document of a translation of an archival item), is it worth the expense to prioritize the digitization of the original item?⁸³

⁸¹ Lorena Gauthereau-Bryson, “Global Linking Through Archival Digitization: The Our Americas Archive Partnership and Spanish, Portuguese, and American Women (1876),” *Microform & Digitization Review* 40, no. 1 (March 2011): 25, doi:10.1515/mdr.2011.003.

⁸² Bantin and Agne, “Digitizing for Value,” 246.

⁸³ Pauline Joseph, Michael Gregg, and Sally May, “Digitisation of the WA Welcome Wall Collection: A Case Study,” *Archives & Manuscripts* 41, no. 3 (November 2013): 191, doi:10.1080/01576895.2013.829752.

Other topics surrounding digitization include patron-added metadata to digitized materials,⁸⁴ user access and need of digitized materials,⁸⁵ how digitization has altered research,⁸⁶ and digitization best practices.⁸⁷ Articles by Emily Monks-Leeson⁸⁸ and Lorena Gauthereau-Bryson⁸⁹ reveal how the impact from digitization projects can affect scholarship. Beyond this point, an exploration of the arbitrary choices made by archivists when deciding which materials will receive a digital surrogate and be made publically available online, as well as those materials that are initially rejected, remains an untouched research subject that is in need of further investigation.

Most of the studies previously discussed only viewed this choice of what archival materials to digitize, if it was mentioned at all, as a small byproduct of a study that was pursuing different research questions. This study is designed to address one question, which is the methodology or, in essence, the choices that are pursued during the selection process of which archival materials will have a digital surrogate created for them. It is hoped that this research will aid in helping those individuals who are setting up digitization programs to be

⁸⁴ Pamela H. Mayer, "Like a Box of Chocolates: A Case Study of User-Contributed Content at Footnote," *American Archivist* 76, no. 1 (Spring/Summer 2013): 19–46.

⁸⁵ Donghee Sinn, "Impact of Digital Archival Collections on Historical Research," *Journal of the American Society for Information Science & Technology* 63, no. 8 (August 2012): 1521–37, doi:10.1002/asi.22650; Jody L. DeRidder and Kathryn G. Matheny, "What Do Researchers Need? Feedback On Use of Online Primary Source Materials," *D-Lib Magazine* 20, no. 7/8 (July 2014), doi:10.1045/july2014-deridder; Gobinda Chowdhury, "From Digital Libraries to Digital Preservation Research: The Importance of Users and Context," *Journal of Documentation* 66, no. 2 (2010): 207–23, doi:http://dx.doi.org/10.1108/00220411011023625.

⁸⁶ Hirtle, "The Impact of Digitization on Special Collections in Libraries."

⁸⁷ Claudia A. Perry, "Education for Digitization: How Do We Prepare?," *Journal of Academic Librarianship* 31, no. 6 (November 2005): 523–32.

⁸⁸ Monks-Leeson, "Archives on the Internet."

⁸⁹ Gauthereau-Bryson, "Global Linking Through Archival Digitization."

conscious of the impact that digitization can have on the collective memory of the public by addressing whether or not personal choice occurs and, if it does, how to address the issue.

Chapter 3: Methodology

This exploratory research project, which is the examination of “a new interest” or “when the subject of study itself is relatively new,”⁹⁰ involved four different case studies. Each of these case studies analyzed a repository’s digital finding aids and the corresponding digital surrogates of a collection, as well as any documentation available; thereafter, an interview with the archivist or other person responsible for the digitization program was conducted. The research was conducted to aid the understanding of the nature and frequency of personal choice by archivists in creating digital surrogates.

Because the phenomenon of personal choice is the focus of this research, inductive methodology was determined to be the best way to proceed. Inductive methodology in this instance was defined as “reasoning from observations,”⁹¹ hence the performance of four case studies. A case study, as defined by John Creswell, is “a qualitative approach in which the investigator explores a bounded system (a *case*) or multiple bounded systems (cases) over time, through detailed, in-depth data collection involving *multiple sources of information* (e.g., observations, interviews, audiovisual material, and documents and reports), and reports a case *description* and case-bound themes.”⁹² This method was also chosen for the lack of any

⁹⁰ Earl R. Babbie, *The Basics of Social Research*, 5th ed. (Belmont, CA: Cengage Learning, 2011), 95.

⁹¹ Lynn Silipigni Connaway and Ronald R. Powell, *Basic Research Methods for Librarians*, 5th ed. (Santa Barbara: ABC-CLIO, 2010), 20.

⁹² John W. Creswell, *Qualitative Inquiry Research Design: Choosing among Five Approaches* (Sage Publications,, 2007), 73. Emphasis is Creswell’s.

experimentation involved; that is, the study will be wholly based upon observation of all the factors already present and included in the study.⁹³

To identify each potential case study candidate, I focused on archival repositories closest to St. Louis, MO (where I live), or areas where I have personal contacts. This search involved reviewing repositories in St. Louis, Missouri; Decatur, Illinois; and Virginia Beach, Virginia. This filtering step was initially thought necessary because an on-site inventory of a collection might be required to determine what, if any, items or collections lack a digital surrogate. However, this need to evaluate the physical collection was ultimately deemed unnecessary, as finding further differences between the physical collection and its corresponding online digital surrogates would not greatly enhance this study's research; all that was needed was to find an instance of personal choice influencing digitization selection.

A minimum number of three case study candidates was arrived at by deciding to investigate at least one institution from the three repository sizes as described by the Heritage Preservation institution: smaller institutions (whose total archival collection is less than 1,000 linear feet), medium institutions (whose total archival collection is between 1,000 and 4,999 linear feet), and larger institutions (whose total archival collection is greater than 5,000 linear feet),⁹⁴ with a goal of being able to investigate two institutions of each size. However, due to lack of response from small and medium repositories and time constraints, a total of four case

⁹³ Connaway and Powell, *Basic Research Methods for Librarians*, 80.

⁹⁴ *A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections* (Washington, D. C.: Heritage Preservation, 2005): 17-8, <http://www.heritagepreservation.org/hhi/HHIfull.pdf>.

studies were completed: one small institution, one medium institution, and two large institutions.

Next was a determination of whether or not the institution had digital surrogates available for online public access. This was followed by a general comparison of any finding aids to the total collection of corresponding digital surrogates available online to determine if any discrepancies were present. Such discrepancies might include an online digitized collection that lacks a digital surrogate of an item listed in the finding aid, but that might have differences in the online finding aid and the online digital surrogates. For reference purposes, I noted the similarities and differences between the finding aid (representing the physical collection) and the online collection. This was followed by seeking any documentation of policies posted online via institutional websites concerning digitization that the repository may have. This would be helpful in both the analysis of policies for inclusion in the final report and in the instance that the associated archivist is questioned about the policies if they conflict with practice or are ambiguous and thus need explanation. The best candidates for inclusion into this study were repositories with some digital surrogates from a collection that were presented online, but whose corresponding finding aid indicated an item or items from the collection for which no digital surrogate was found online. A total of thirteen archival institutions were contacted at the outset of the study.

The final step in the process was setting up an interview with the archivist or other knowledgeable personnel involved in the decision-making process of which materials are digitized as well as which digital surrogates are placed on the Web. This was done after a point

of contact was first determined for each potential case study candidate. Once established, each candidate was contacted initially via email for permission to participate in the study (see Appendix A for the contact email). The sent email explained the scope of the study and asked if there was someone at the archives who would be willing to participate in the research. If the individual accepted the invitation, they were sent a copy of the questions to be asked during the interview, and a day and time were set up to meet for an in-person interview. The interview covered the questions (see Appendix B) of the selection of archival materials to digitize, and lasted between 30 and 90 minutes. Each interviewee was the current person overseeing the digital collections. The case study interview process primarily involved open-ended questions concerning what ideas went into the decision process.

To be able to recall information more accurately, I asked each participant if there would be any problems with having the audio of the interview recorded. All participants verbally agreed to allow audio recordings. After the interviews, all audio was completely transcribed to be analyzed. This content analysis was performed on the transcription of the interviews to chart dominant trends based on statements. The chart contained main headings for each category, and all relevant answers were entered under the appropriate heading for each institution.

These headings are listed below.

- Reasons to digitize.
- Digitization protocol.
- Storage of images.
- Digitization of materials.

- Patron access.
- Number of items/images digitized.
- Use of volunteers/student help for selection of materials to digitize.
- Use of volunteers/student help to digitize materials.
- Funding for digitization.
- Other facts.

The data from each case study were altered enough to help each institution and interviewee remain anonymous. The results from the content analysis were then combined to understand how the personal choice of what to digitize was impacted by the interview answers. The next chapter presents the results of the four case studies from the interviews with personnel currently in charge of digitization. Following that, Chapter 6 will present how the information gleaned from the content analysis reveals the use of personal choice to determine which items have derivative digital surrogates created.

Chapter 4: Case Studies

The four case studies are presented anonymously and will be listed as Institution A, Institution B, etc. Anonymity was done to protect the identity of the interviewees as the purpose of this study was only to find evidence of personal choice influencing digitization decisions and was not trying to single out either people or institutions. The presentation within each case study will describe several factors about each institution:

- The general size of the physical archival collection.
- The number of staff within the department.
- The number of online digital surrogates created.
- The source(s) of funding for digitization.
- How decisions are reached regarding what to digitize.
- The types of materials that get digitized.
- Whether or not digital surrogates are publically accessible.

Before proceeding to the synopses of the institutions, a definition should be reiterated for the benefit of the reader. When referring to digitized surrogates, I mean to reference each individual digital file created and stored on a computer medium, such as a hard disc, CD-ROM, flash drive, server, etc.⁹⁵

4.1 *Institution A*

⁹⁵ Please refer back to footnote 1 for a fuller definition of 'digital surrogate.'

Institution A consists of an archive with more than 300 collections and surpasses 5,160 linear feet of materials. Three full-time staff are employed within the archival department. The interviewee was unable to determine the total number of individual surrogates created from digitization but did state that more than 16,000 items from the archives have been digitized. However, the interviewee stated that items may contain multiple surrogates, such as a book with 100 pages; in this instance the book is counted as a single digital surrogate rather than 100 digital surrogates.

Funding for digitization projects is derived from general departmental monies, patron requests, and grant awards. For patron requests, any costs are determined on a case-by-case basis and will be passed along to the patron. For example, if the patron requires that the digital surrogate be placed on a DVD and shipped to them, the archives would charge the patron for the cost of the DVD and postage. The interviewee did indicate that a specific policy addressing such costs does not exist.

The determination of which items Institution A digitizes for access is multifold. Patrons may visit the archives and request to have a digital surrogate of specific items, or patrons may contact the archives from remote locations and make a request for specific materials. These requests are granted barring any access restrictions on the materials. Patrons from other departments within Institution A have also requested the creation of digital surrogates. This included an instance in which a physical exhibit was developed to celebrate a famous personality who once attended Institute A. The archives, which housed a collection of materials

from and about this celebrity, was asked to develop an online exhibit that used only some of the archival materials within the collection.

The interviewee also cited preservation of archival materials that receive a high volume of usage as another factor that may determine when an item is digitized. As stated during our conversation, they decided to digitize some fragile, yet significant, diaries from the 1800s due to the poor conditions of the materials. This was also the case with a different collection of archival materials. The interviewee stated, “we had . . . maps from the 1800s, also. And those are getting some use, and they would just crumble. We would have the little shreds of paper on the floor. And so that got worked into another grant with the main library here and the [local museum].”

Successful grant projects also play a role in determining materials to be digitized; for example, Institution A collaborated with a state grant-funded initiative to digitize materials within their archives relating to the Civil War. With this specific grant project, the determination of what was to be digitized was partly made by another institution that was the project leader for the grant. Although not everything within a whole collection was digitized, all digital surrogates that have been created are available to the public for viewing.

The kinds of materials that are digitized include printed and manuscript items, photographs, audio, video, and over-sized printed materials. The interviewee noted that the over-sized items are contracted to a third-party for digitization as Institution A lacks the means to digitize such items.

When Institution A decides to digitize non-grant and non-patron-requested items, it is driven mainly by “patron request,” as stated during the interview. The interviewee noted that they and the staff convene to determine what is receiving the most patron demand for items not already digitized. Those items are given priority for digitization over all other materials within the archives that have not been digitized already. However, the interviewee stated that the majority of their archives is post-1923 and under copyright restrictions.

Patron access to the digital surrogates from the archives is provided via its website. Low-resolution⁹⁶ images for printed materials are available for public viewing via the Archon software platform,⁹⁷ and the original high-resolution images are available upon request. According to the interviewee and as noted earlier, a “process and handling” fee is charged to the patron to send them any requested high-resolution versions of an image.

Institution A lacks a digital assets management platform for easy storage and retrieval of digitized surrogates. All digitized surrogates are currently stored locally on onsite computers. Some digitized textual surrogates have been uploaded to HathiTrust.com and GoogleBooks.com, third-party organizations that host print materials that have been digitized.

⁹⁶ According to the National Archives and Records Administration, spatial resolution is how “finely or widely spaced the individual pixels are from each other.” In common parlance, spatial resolution is shortened to simply ‘resolution’ which is the term I will use. The greater the resolution, the more detail an image, audio, video, etc. digital file contains through the increase of pixels within a given area. Steven Puglia, Jeffrey Reed, and Erin Rhodes, *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images*, National Archives and Records Administration, June 2004. <http://www.archives.gov/preservation/technical/guidelines.pdf>

⁹⁷ Archon was an archival software platform originally developed by the University of Illinois (<http://www.archon.org/>). Support for the software was discontinued in 2014 as the developers merged with the developers of Archivist’s Toolkit, another archival software platform, to create a new archival software platform called ArchivesSpace (<http://www.archivesspace.org/>).

4.2 *Institution B*

The archives at Institution B consists of 206 collections, which comprise about 1,987 linear feet. The digitized content consists of approximately 12,315 digital surrogates. At Institution B, only two staff members are involved in the digitization process.

To date, only photographs, textual materials, and audio have been digitized by the archives at Institution B. They have not digitized any motion-picture materials. All digitization is done in-house by either staff or student assistants. The only funding source for the archives at Institution B has been received from the departmental budget. Institution B has not received any grants nor does it charge patrons for access to high-resolution copies of digital surrogates.

Institution B does not have a specific written protocol to determine the prioritization of materials for digitization, but does have a digitization policy that addresses information about the digitization program and patrons, as well as providing guidelines about how to select materials for digitization. Listed below are the criteria outlined in the policy that materials must meet to be considered for digitization:

- The institution has the right to digitize the material.
- Does not duplicate existing projects.
- Increase the informational and educational needs of the community.
- Materials can be accessed without a fee.
- The material can be physically digitized.
- The materials form a “cohesive whole” (parts of a collection may not be digitized).

- The materials will not be damaged.
- The technology, funding, and staff knowledge are present to digitize the materials.
- The digitized materials will have the necessary support to be sustained over time.
- Significant to the institution, an upcoming exhibit, or is valuable.
- Adds to the completeness of already digitized collections.

The interviewee did provide a list of collections that was created in 2012, in order of priority, that the archives hopes to digitize. However, according to the interviewee, the determination of what to digitize is first and foremost driven by institutional inter-departmental collaboration or student projects. When resources are not being utilized for specific student or inter-departmental projects, website search statistics were cited as also being used to gain insight into what topics and/or archival items patrons are seeking. These data directly influence the interviewee's determination of what to digitize. The interviewee did not say how the use of such statistics fits within the digitization policy or the priority list.

At this archives, all items must also be evaluated according to ownership, equipment requirements, and condition before being considered for digitization. In addition, any items deemed to have sensitive, identifiable information are rejected for digitization for public viewing. This information was received during the interview and corroborates their digitization policy. Unlike the situation with Institution A, not all digital surrogates are available for public access. An unknown number are kept away from the public, not due to any restrictions but

rather the interviewee expressed a desire to make the images part of an online exhibit before making them available to the public.

Patron access to digital items is via the archives' website running the CONTENTdm platform.⁹⁸ From here, patrons are able to view or retrieve full-resolution scans of the digital surrogates without any access restriction. All files for digital surrogates are stored locally on institutional servers.

Not all digital surrogates are available to the public. During the interview a student digitization project was brought up. The student only desired to create an exhibit that focused on a specific aspect of the collection; however, most or possibly all the collection was digitized. When asked about the digital surrogates that had been created but are not available to the public, it was stated, "we'll do something with them eventually. They are scanned. They are digitized, but they are not online . . . [We are] trying to figure out how they could become their own collection."

4.3 *Institution C*

The archives at Institution C consists of 8,360 linear feet of materials within nineteen collections. The archives has two full-time staff members. However, they also utilize students to aid in digitizing materials.

⁹⁸ CONTENTdm is a third-party digital asset management software system provided by the Online Computer Library Center (OCLC). <https://www.oclc.org/contentdm.en.html>

The materials that the archives have digitized are textual materials (both manuscripts and printed), photographs, and audio. All digitization is done at the institution by staff and student assistants.

The funding to digitize the archives within Institution C has only been through general department funds. Student and inter-departmental projects have also occurred, but the funding is still derived from the archive's parent institution. At the time of the interview, Institution C had not received any funding from grants or from patron requests.

With its digitization procedures in place, Institution C's archives has been operating without an approved digitization policy. A written policy has been created, but it has never been approved by the institution's administration.

Patron access to digitized surrogates is through a statewide consortium website that relies on CONTENTdm to allow management and access to each participating institutional archive. Patrons are able to access image resolution sizes up to 1,000 pixels per side. Larger or full-resolution images have to be requested from the institution's archives. All digital surrogates created are stored offsite through an agreement with a university within the United States to host and preserve the original digital surrogates.

The interviewee at Institution C provided several reasons for digitizing. Among them included institutional memory through the digitization of many of its publications. Because the institution owns the copyright on such items, and barring any other laws that may restrict public access, it may place such materials online without restriction. The interviewee also mentioned collaboration on projects with other departments within the institution as a reason

to digitize; one such project involved documenting the institution's botanical plants. This was a collaboration between the archives and two other departments.

The interviewee did indicate that, to date, all items that have a digital surrogate are available online for public access. They also noted that although all digital surrogates are available, that did not mean that all materials within a collection had been digitized.

Exposure of the collections was another justification for digitizing archival materials. The interviewee stated that they hoped to generate interest about the collections through the digitization of some of the materials from said collections, as well as through their inclusion to the online repository for public access.

4.4 *Institution D*

Institution D's archives consists of a single collection where all materials are not separated into distinct collections. Within the archives are about 2,000 items that comprise approximately 60 linear feet. Their digital content currently stands at about 5,000 digital surrogates. The interviewee was unable to give precise numbers for either the total number of items within the archives or for the total amount of digital items created. Although 2,203 digital surrogates are listed as being created, this number comprises items such as pamphlets, which are counted as a single file with multiple images. This made determining the exact number of individual digital surrogates more difficult.

The institution has two full-time staff members, and does utilize student assistance to digitize its archival materials. However, volunteers have not participated in digitizing any

materials. Funding for the digitization of materials is derived from general institutional funds, patron requests for copies of the digital surrogates, and grant awards.

The archives does not have a written policy to follow regarding digitization, so the procedure by which to determine what to digitize is very broad. To summarize their digitization philosophy as expressed by the interviewee, they intend to scan everything within the institution's collections that relate to the institution. A copy of each digital surrogate is kept locally on a server. The interviewee stated that not all digital surrogates are available online due to one of two reasons: either the digital surrogate lacked metadata associated with it before being uploaded or the digital surrogate was not related to the history of the institution.

Among the materials scanned by the archives are photographs, manuscript items, and three-dimensional objects (such as plates, paintings, etc.). No video or audio materials have been digitized to date.

Institution D's digital surrogates are hosted through a third-party website – PastPerfect.⁹⁹ Through this website, patrons are able to access low-resolution, watermarked versions of the digital surrogates. High-resolution copies of the digital surrogates are available for a fee from the institution upon request.

Based on these four completed case studies, some observations can be made regarding the processes used to determine which archival materials will be digitized for public access. The impact of these procedures upon collective memory and the aforementioned digitization of materials will be discussed further in the next chapter.

⁹⁹ PastPerfect museum software for content management. <http://museumsoftware.com/pponline.html>

Chapter 5: Analysis

Presented here is an analysis of the information acquired from the four case studies.

The analysis will look for any trends the four institutions share as well as noting any information that takes a minority position. These trends are summarized in Table 1 about reasons given for digitization which is followed by a deeper look into each of the reasons listed in the table. After discussing the reasons for digitization, other areas not brought up by the interviewees will be investigated: donations and financial influence. The former was cited by each interviewee as not influential in their decision about what to digitize while the latter was not mentioned; that said, this researcher believes these matters are salient to the discussion of what influences personal choice when digitizing materials. First, a broad overview of the analysis of the data is presented.

5.1 *Overview*

The four institutions analyzed for this study present a cross section of archival repository sizes, both in terms of collection size and number of staff. A comparison of the findings from each of the institutions reveals some similarities as well as specific differences. Each of the institutions presents different mixed methods of approaching the decision process for determining which archival materials are to receive digital surrogates. Table 1 indicates the identified categories, that is, the reasons each interviewee stated for determining how archival materials are chosen for digitization. Arbitrary colors were used in the table to help highlight each answer. The colors do not correspond to any other meanings and are only present to aid the reader.

Table 1. Comparison of the reasons for digitizing per institution.

Reasons	Institution			
	A	B	C	D
Collaboration	Yes	Yes	Yes	N/A
Patron requests	Yes	Yes	No	Yes
Student projects	No	Yes	Yes	No
Exposure	N/A	N/A	Yes	N/A
Administrative requests	N/A	N/A	Yes	N/A
Preservation	Yes	N/A	N/A	N/A
Donations	No	No	No	No
Financial	N/A	N/A	N/A	N/A

Color Key:	Yes	No	N/A	
-------------------	------------	-----------	------------	--

Upon analysis of the interview transcriptions, distinct reasons for digitization brought up by the interviewees became more apparent. In some instances, a distinct answer was given by the interviewee that set itself apart from other answers, such as ‘preservation’ and ‘exposure.’ This led to other institutions’ answer(s) being classified as ‘N/A’ (Not Applicable).

Overall, collaboration and patron requests were cited as the most popular reasons to digitize materials. This was followed by student projects that were cited as a reason to digitize materials at two of the institutions. Cited only once were the following reasons given for digitizing: requests from an institution’s administration, preservation, exposure of a collection and/or the institution. Each reason will be addressed in the order listed in the table, from top to bottom.

5.2 Collaboration

Amongst the most popular reason for digitization of materials was collaboration with other groups, usually from within the archives' parent institution, but, in one case, with an outside entity (Institution C). Of the four institutions, three interviewees stated this as a reason for digitizing archival materials. This method of choosing removes the need for the person or persons within the archives to be responsible for what materials from their collection to digitize. Instead, collaboration with other parties allows the archives to mitigate such responsibility for the decision, and has the added benefit of developing collegial bonds and defraying some of the costs involved in digitizing, which was something noted by Ken Middleton.¹⁰⁰

The interviewee at Institution B mentioned that collaboration was part of the impetus for creating a formal digitization program. After working with another department to digitize a specific collection, the archives developed formal procedures for digitization. As the interviewee stated, the benefit of this first digitization project was that it “fosters collaboration and enhances the library’s role on [locally] or in the community.”

As detailed earlier, the interviewee at Institution C mentioned a unique situation whereby they collaborated with an outside institution on a digitization project. Institution C digitized a Civil War diary that neither belonged to them nor was part of a collaborative exhibit, nor did Institution C have a collection that pertained to the diary. This discovery was made after the interview and thus was not addressed by the interviewee. This was the only representation

¹⁰⁰ Ken Middleton, “Collaborative Digitization Programs: A Multifaceted Approach to Sustainability,” *Library Hi Tech* 23, no. 2 (2005): 146-50.

of an institution collaborating with a third party to digitize their content and provide internet access to the third-party's digital surrogates.

Institution D's archives, the smallest of the four, presented the most unique scenario in that they were the only archives that did not cooperate with other parties. Cooperation with other institutions was not a subject broached by the interviewee. The lack of volunteering any information does not automatically negate the possibility of collaborating with other institutions.

5.3 *Patron requests*

Patron requests also received three citations as a reason for digitization. This method of deciding what to digitize is very similar in principle to collaborating with other institutions in that the decision mechanism of what to digitize is removed from the archival staff and is being dictated by an outside influence.

Institution A cited patron requests as the primary determiner of digitization. Patron requests are noted, and if the item had not been digitized, the institution would proceed to do so when not already working on a specific project.

Institution D acknowledged that they process patron requests for digital surrogates of items not already digitized from their collection. They did not say whether or not the requests had to be for materials relating to the institution's own history. The interviewee also did not state how much importance was given to patron requests.

Only the archives at institution C actively turns down patron requests for the digitization of archival materials. The interviewee stated, “. . . I don’t scan anything for anybody, because there’s no policy. There’s no procedure. I can’t even tell them, ‘I need to charge you this much to recover the cost of the labor.’”

5.4 *Student projects*

Institutions B and C explicitly cited digitization being driven by student projects. In the case of Institution B, one of the early digitization projects was undertaken by a student who was perusing the archival materials and found a miscellaneous collection mostly consisting of correspondences with a 19th-century editor. This collection is notable for the many correspondences with prominent figures of the day. However, several of the collection’s contents were not digitized. The reason for the omission of some of the collection’s content was described by the interviewee as not pertinent to the student’s project. In addition, other items from the collection were digitized but not placed online and instead reside on local media. An effort to remove this gap in access through digitization of the remaining content from the collection was not completed by either the student or the institution. After reviewing a description of the non-digitized materials as listed in the finding aid, it seems the items may not be relevant to the collection they are located with, but may be more relevant to other collections outside of the institution. Without a digital surrogate, it is difficult to determine whether or not these materials would be beneficial in any context.

Institution C also acknowledged that student projects were a driver of digitization. The interviewee noted that the student was from a different department but had worked with

Omeka content management software.¹⁰¹ As the interviewee stated, “the student and I have created, using the content that I am responsible for, as our first project, a tribute to the . . . anniversary of the opening of this campus.”

5.5 *Exposure*

The idea of digitizing archival materials for exposure is a very broad concept with which all of the four institutions in this study would likely agree. Although only Institution C was explicit about exposing collections as a reason for digitizing, it can be said that all digital surrogates that are placed online for public access are likely to benefit their institutional patrons with public interest in their collections. Because many institutions are placing their digital surrogates on the World Wide Web for patron access, it can be safely assumed that their collections will gain exposure via the harvesting of web content by internet search engines. If an institution did not want exposure of their collections, making the digital surrogates available for mass public access would be the wrong way to proceed.

For this reason, it is not known whether ‘exposure’ should be considered a valid category for wanting to digitize archival materials. Nonetheless, because it was specifically stated by the interviewee, it was decided to include it here.

5.6 *Administrative requests*

Administrative requests was mentioned by Institution C as a reason for digitizing items. This reason refers to an instance whereby the institution’s administration had a hand in the

¹⁰¹ Omeka is a 3rd-party online content management software platform for curating digital collections. <http://omeka.org/>

decision to digitize part of a collection. As the interviewee stated, “To be respectful and responsive to the family, to the donors, our administration decided we would scan that small portion. So we’ve created . . . what I would think of as a mini-digital collection.” Here, the involvement of third parties is influencing the digitization of materials. In this situation, the archivist has had the decision of what to digitize removed.

5.7 Preservation

One institution, Institution A, also cited preservation as a reason to digitize. As mentioned during the description of Institution A in Chapter 4, the interviewee explained that preservation was a consideration for reducing wear and tear on heavily used archival materials. None of the other interviewees in any way mentioned reducing patron access as a reason to digitize either specific items or an entire collection.

5.8 Donations

One question asked of each of the participants was whether or not they accepted donations with the main goal of wanting to digitize the collection. It was thought that this might have been a way in which archives might use personal judgment to decide what materials gain greater public recognition and access. All four institutions said that digitization has not been in the past or present a factor in accepting donations.

5.9 Financial

One driver of digitization that was not brought up by any of the four interviewees was money. The impact of money as a driver of digitization is implicit yet very tangible. Three

specific instances were identified where finances implicitly drove digitization: grants, donations, and patron requests.

Grants and monetary donations can be lumped into a single group, as the similarity of the two can be seen by looking at how each interacts with the recipients. In the case of both grants and donations, there is an impetus to define digitization projects that will be able to draw interest from a donor (whether in the form of a grant or direct financial assistance coming from either a single person or a group) and be able to win greater support than other institutions that may be vying for the same funding.¹⁰²

Conversely, digitization projects that do not provide a tangible return on investment from grantors and donors may be deemed a failure. These failures could negatively impact funding for digitization projects on the whole. In addition, projects deemed a waste of money can cause a chilling effect on similar (in this case, of digitization projects) grant-funded endeavors.¹⁰³

Patron requests can also implicitly play a role in determining what to digitize. As noted with Institutions A and D, patrons are potentially charged for access to digital surrogates. However, it should be noted again that Institution A said they would charge patrons a fee only if they themselves would incur a direct fee (such as purchasing media to save the digital surrogate to, paying of postage, etc.). Institution D did not specify the cost involved for patrons

¹⁰² Having personally participated in assisting with applying for a grant to digitize a collection, and having written and won several personal grants, I can attest to the drain on time and resources that writing a proposal for fiduciary gain can be.

¹⁰³ Joanna Kempner. "The Chilling Effect: How Do Researchers React to Controversy?" *PLoS Medicine* 5, no. 11 (November 2008): e222.

to access full-resolution copies of digital surrogates. It could be a simple matter of control of the intellectual property rights of the digital surrogates. And no implication is meant that the archives are necessarily actively trying to profit from any income received from licensing digital items for third-party use as the funds may simply go back into the operating expenses for the archives.

Unfortunately, information about how money influences digitization, or social research, appears to be wanting. Perhaps this area of study will be picked up by someone else.

With a look at the reasons for digitization, whether explicit or implicit, the discussion should turn toward what was not been digitized. The next chapter will look at the subject of what was left out of the digitization process, as well as policies in place guiding digitization and the involvement of personal choice in such selections.

Chapter 6: Discussion

With a look at the reasons stated by each interviewee regarding the selection process of digitization, attention can be turned toward the impact of those reasons. This will mainly focus on what archival materials were implied not to be digitized. But first, a look needs to be done at the policies guiding digitization at each of the institutions.

6.1 *Policies*

Of the four archives, only Institution B has a formal written policy regarding the digitization of archival materials. The informal policies that Institutions A, C, and D are utilizing appear to be based more on habitual practices developed over time than policies that have been researched, written, and approved by any administration. The one exception was Institution C, which has a written policy that has never been reviewed and approved by the institution's administration, although it has been submitted; thus, it is not being counted as a formal policy for this thesis. The interviewee said they had developed the written policy for digitization based on information received from the Northeast Document Conservation Center's guidelines.¹⁰⁴ Although Institution C does not have a formal policy, the interviewee stated the following rationale regarding the criteria for selecting materials:

. . . what collections do we have? And what collections, if I'm the person who is in position to make the call as to . . . who actually uses it; is likely to use it; how does this fit the agenda of the institution I serve, the library? For us, it just needs to be a process that I make a decision. I need this stuff scanned, and the people who accomplish that carry it out. It's a question of practicality. At a different

¹⁰⁴ The Northeast Document Conservation Center offers guidelines for the complete digitization process from the initial decision of proceeding with digitizing to the preservation of digital items. - <https://www.nedcc.org/preservation101/session-7/7digitization>. Accessed November 8, 2015

library with different resources, human and technological, I think you can afford to be more idealistic. I still think about the criteria, but the main criteria [we are] concerned with are: do we own it? Do we own both elements? Do we own it physically and [do] we own the intellectual property rights? Beyond that, probably how much work is this going to be for the technician who does the scanning and the metadata librarian? And in the mix of materials that I hold, and the agenda of the library at the moment, how does this fit in?

As can be noted in the quotation, the interviewee at first describes a situation whereby they simply determine the criteria for what to digitize. This thought then morphs into more subjective ideas about intellectual rights, technical capabilities, and the amount of staff time needed to implement a project.

With regard to the approval and implementation of the draft policy that was developed, the interviewee at Institution C stated, “I was very principled and idealistic about this when we started. And what I have found is that in addition to the criteria, we set up a process of consideration involving way too many [people]. And my intent as the curator essentially got lost. We were focusing too much on the conversation. And [when] we met in person . . . it proved basically impossible to get everybody together; and everybody included administrators. It proved so clunky and unworkable that I basically subverted it because we weren’t getting anywhere.”

Institution D has a single, informal policy of digitizing everything pertaining to the institution. Once completed, all archival material, and many of the institution’s three-dimensional items (plates, cups, clothing, etc.) will be made available via its website as digital surrogates.

6.2 *What's not digitized*

Another of the questions that was asked of each interviewee was whether or not there had been a situation whereby only part of a collection has been digitized. Institutions A, B, and C said that there have been projects whereby only a part of a collection was digitized for public access.

As detailed previously, the archives at Institution A mentioned two specific instances where materials from a collection were not digitized. The interviewee cited an instance of collaboration with another department to create an online exhibit to coincide with a physical exhibit. In this circumstance only specific items were needed to be digitized from the collection that were relevant to the physical exhibit. Another such instance mentioned earlier was a grant-funded collaboration whereby Institution A digitized some materials from its collection relating to the Civil War. Both examples illustrate a situation where only select archival materials were digitized, but the choice of which materials to digitize was partly dictated by a third party. This does herald a different question about whether or not a participant should try to advocate for the digitization of an entire collection, regardless of whether or not all materials will be used as part of a virtual exhibit. As noted by Barbara Craig, Susan Crane, and Ryan Davis earlier, access to primary source materials helps shape collective memory.¹⁰⁵ The interviewees at Institutions

¹⁰⁵ Craig, "Selected Themes in the Literature on Memory and Their Pertinence to Archives," 276–89. Crane, "Memory, Distortion, and History," 44. Davis, "Caught in Franco's Web," 90.

B and C both noted student projects that utilized only specific parts of an archival collection in order to develop a specific online exhibit.

Institution D withheld digital surrogates, stating that although they have scanned materials in their archives that do not pertain to the history of the institution, and their account with PastPerfect would accommodate the extra digital surrogates, it was not currently their intention to pursue uploading those digital surrogates. Their goal is to only make all institutionally related archival materials available for public access. All other materials within the collection are to remain without public access to a digital surrogate. The interviewee stated that this unwritten policy of not providing online all digital surrogates could change one day. Also, although the digital surrogates that do exist but do not directly relate to the institution are sitting on a local server, the institution does accept requests to digitize archival materials not relating to the institution.

This lack of digitization is an example of controlling a specific narrative around a subject. In the example of Institution A's online exhibit centering on a celebrity that tied in with a physical exhibit, the institution controlled how people generally viewed the exhibition subject. This harkens back to the example presented regarding events as told in *History Wars* surrounding the Smithsonian's attempt to present a more nuanced and accurate depiction of the events occurring at the end of World War II.¹⁰⁶ A major difference between the two situations was that Institution A was a willing participant in developing a specific narrative regarding the celebrity whereas the Smithsonian was coerced into altering its exhibit to

¹⁰⁶ Linenthal and Engelhardt, *History Wars*.

promote a specific narrative. Without all the information on an event or person readily at the public's disposal, critical discussions to shape collective memory cannot take place.¹⁰⁷

Institution D also presents two different means of controlling public access to digital surrogates. Patrons are required to contact the institution to gain access to a full-resolution version of the digital item, and this may require a financial transaction (at the institution's discretion).

When looking into whether or not any of the four interviewees noted instances of withholding digital surrogates from the public, the person at Institution A explicitly said, "we don't usually withhold anything. And we haven't taken anything down." Institution C complimented this idea of a desire for not withholding any digital surrogates or removing them from public access. When asked this question, the interviewee at Institution C said, "So far, no." In all three of these cases, there seemed to be a hint of reservation about committing to a definitive 'no' answer. The interviewee at Institution A included the words 'we don't usually withhold' even though, to date, nothing has been removed or withheld from public access. Likewise, the interviewee at Institution C used the qualifying phrase 'so far' before saying 'no' despite not having removed or withheld any digital surrogates.

This left only Institutions B and D, both of which have withheld digital surrogates from the public, but for very different reasons. As stated earlier, Institution D withholds digital surrogates that do not directly relate to the institution's history. On the other hand, Institution B, which has withheld digital surrogates, does hope to provide them to the public once the

¹⁰⁷ Kerwin, "On the Emergence of 'Memory'," 127.

personnel have decided how to present them online. The interviewee at Institution B did state that with some digital surrogates of oral histories the institution had done, there was a possibility of having to remove the digital surrogates if a request was made.

6.3 *Evidence of personal choice*

This study did reveal that in the case of these four institutions, the personal influence of those digitizing materials did play a role in determining what patrons have access to in an online environment. In the case of Institutions B and C, this influence can be attributed to the desire to create virtual exhibits. This does hearken to a question in Chapter 1 about whether or not archivists, or their assistants (i.e., student workers, volunteers, etc.), can be considered experts on the collections with which they work. This can be especially pertinent when pieces of disparate collections are being digitized to create a single, new exhibit, which in turn might develop into a new narrative.

In the case of Institution A, parts of a collection were digitized for an online exhibit to coincide with a physical exhibit about a celebrity associated with the institution. It was the concerted effort of those involved with the exhibits to pick and choose from the collection to decide what to utilize, including which materials would get a digital surrogate.

Institution D made a specific decision to only digitize and make publically available digital surrogates of collection materials that relate to the history of the institution. This decision leaves all other materials not relating to the institution, yet within the archival collection, more hidden than they otherwise would be if the institution did digitize and make available all their materials.

It should also be stated that the instances of personal choice influencing the digitization of materials appears to be in isolated instances and does not appear to be pervasive among the participating institutions. All participants stated that they try to adhere to policies, whether written or unwritten, when making decisions about digitization. In some cases metrics were utilized to determine what to digitize, in other cases projects were determined by student projects or collaboration with other institutional departments.

6.4 Impact of the findings

It is hoped that this paper brings attention to the influences that personal choice can have on the digitization of archival materials. Such choices can impact public access to digital surrogates of primary source items, which in turn can influence collective memory. It is also hoped that these case studies will spur further research to determine the extent of personal choice when deciding what primary source materials get digitized. As shown in Chapter 2, much has been written about the digitization of archival materials. But the current research here was conducted to focus on one facet of digitization that has received very little attention.

The development of policies is one step to minimizing the effects of personal choice on the selection of archival materials to digitize. Policies can help guide the selection process by becoming a proxy in the decision-making process for an individual. Diligence and adherence to such developed policies may aid in mitigating personal choice as a factor.

Using data analysis, such as done by Institutions B and C, to help determine which collections bring in the most inquiries can help remove personal biases from the selection process. Data analysis can also be utilized to create a priority list from which to work.

Another possible way to minimize the effects of personal choice is to involve multiple people to aid in the consideration of which materials and collections to digitize. Creating committees that involve people of different subject specialties and ideological backgrounds is one solution. Such committees can help mitigate any of the biases that each person carries with them by diffusing the pool of personal choice among multiple influences.

Similar to the development of committees to reduce personal influence when determining which collections to digitize, an area that is also touched upon in the case study results is creating new exhibits through the partial digitization of collections and thus could be beneficial to creating new narratives. But these narratives need to be carefully considered and should be vetted by multiple personnel, again similar to the selection of what materials to digitize; these personnel should include subject specialists in the collection being represented (e.g., a musicologist and trumpet playing professor in the Jazz Music Department for a collection on the iconic 20th-century jazz musician Miles Davis) and individuals with varying points of view. Unlike defining a new narrative to shape collective memory through writing a research paper or a journal article, which will likely receive less exposure through the restrictions often accompanying those forms of communication and cannot be as easily disputed in print, creating a new narrative that is freely and widely accessible should be done with greater caution.

6.5 *Study limitations*

This is a project that, in hindsight, may have been conducted in a different manner. One of the greatest limitations was the total number of case studies completed. The limited data set

of this research is too small to represent the trends regarding the influences driving today's digitization of archival materials. However, it is not believed that more case studies would have been beneficial. This research established that personal choice does play a role in the decision making, but it has not established the degree to which personal choice occurs nor how often it occurs. The degree to which personal choice occurs and how often it occurs might be better addressed with larger data sets. For example, a general survey sent to multiple institutions that are digitizing archival materials would likely have provided greater insight into and a much broader scope of trends when addressing institutions and their decisions about the digitization of their archival materials. Along this line of reasoning, it is not believed that the size of the archival institution, based on the amount of linear feet of materials housed by the archives, would impact the results of this study in any meaningful way. However, a larger data set may establish possible links between the size of the archives and the influence of personal choice on digitization.

In addition, by providing a list of reasons for digitizing materials, study participants would be able to read a list of answers that would aid in memory recall when answering a question about the reason to digitize. Also, such memory recall might instigate further expansion of the answers by the study participant when addressing the specific reasons they digitize.

During this research, additional questions were not asked of any of the participants so as to create continuity among the answers received by the interviewees. This was a limitation as it meant that answers gleaned from interviewee could not be applied to the next

interviewee as it would introduce an incomplete data set. As it was, upon analysis of the data, there are noted instances whereby a single interviewee brought up an instance of a specific influence upon the digitization decision process; for example, the interviewee at Institution A mentioning preservation as a reason to digitize.

Perhaps a general survey being sent out to multiple archival institutions asking the reasons why they digitize would have helped establish a greater set of questions with specific reasons mentioned for digitization. This would possibly have jogged the recall of the interviewees as to their own selection process, however, this would not have aided in establishing that personal choice occurred in deciding what materials to digitize.

6.6 Future research

Although these four case studies have demonstrated that personal choice does play a role in the selection of materials that are digitized, and thus impacts collective memory, more research should be done to understand the extent to which such personal choices play a role in the selection of archival materials to digitize.

As mentioned in the previous section, two questions that need to be addressed are: 1) How often does personal choice occur in the selection process of digitizing materials? 2) To what degree does personal choice influence the selection process?

Another question that could be addressed by further research is the influence of money on digitization. For example, how much influence do patron requests, possibly involving financially lucrative access, have on creating digital surrogates? If archives are able to

quantitatively view statistics on patron traffic, how is this information utilized and what are the outcomes from such information? How much digitization is driven by financial boons such as grants and financial donations? What affect could this have on the digitization of archival collections or parts of collections? Are monies influencing collective memory? If so, what is the impact?

Although this study had limitations, with all four institutions revealing the presence of personal choices influencing selections there should be a greater light shined on this aspect of digitization. Beyond investigating how pervasive this influence is, it would additionally be interesting to reveal whether or not different institutional types play different roles in the selection process. Does the influence of personal choice play a part among corporate archives, governmental archives, volunteer archives, etc. I suspect that it is the case that many, if not most all archives, can point to instances where personal choice of the archivist(s) has dictated the digitization of materials.

This discussion of the evidence for personal choice influencing the selection of archival materials for online public access has demonstrated that such influence occurs. Again, it should be reiterated that it was not meant to establish the degree or the amount to which this personal choice occurs.

Chapter 7: Conclusion

This research aimed to unravel the issue of whether or not personal choice affects which archival materials get digitized into digital surrogates and are made available to the public via the internet, and which materials do not get such treatment. This was done through the examination of four repositories to develop a case study of each. Multiple factors were looked at, including whether or not written and approved digitization policies were present at each institution and by what means digital surrogates were chosen to be created from among the archival materials at each institution. From this exploratory project, it has been determined that personal choice does play a role in the decision-making process of creating digital collections and presenting them online. With all four institutions, an instance of a person directly influencing the digitization process could be revealed. Although three of the institutions lack an implemented written policy regarding the decision process for digitization projects (only Institution B has a written and approved policy), Institutions A and C do follow informal procedures to prevent or ameliorate personal choice entering the decision-making process. Institution D, because of their simple protocol to digitize everything relating to its own history, has not removed personal choice entirely from playing a role in the decision process. It has been noted that a choice not to include archival materials that do not pertain to the institution's history is part of an unwritten policy and these non-institutionally related materials will not be made available for public online access.

Personal choice is not always a negative influence on collective memory and should not be curtailed or quashed. As discussed previously, new collective memory narratives can be

derived when these choices are thoughtfully, with expert participation, used to combine materials from disparate collections. However, these choices to deliberately leave parts of a collection without a digital surrogate that are not initially deemed pertinent to collective memory reduce the public's ability to judge the collective memory narratives that have been developed. Plus, leaving materials without a digital surrogate does not allow the refutation of collective memory narratives by the public.

How individual societies and groups color and shape their sense of history, their collective memory, is dependent upon what sources of information are readily and easily available to them. Although some factors that are beyond an archivist's control can dictate what original source materials get greater access through digitization and online placement, there are those times when choices can be individually pursued. It is those moments of individual decision making that can have an influence on cultural memory. As Janine Solberg remarked concerning the influence of the choices individuals make with regard to digitizing, "[they are] actively cultivating an awareness and dialogue around digital tools, and not treating digital search and discovery activities as transparent, neutral, or inconsequential to the acts of invention and interpretation that lie at the heart of rhetorical histories."¹⁰⁸

¹⁰⁸ Janine Solberg, "Googling the Archive: Digital Tools and the Practice of History," *Advances in the History of Rhetoric* 15, no. 1 (January 2012): 60, doi:10.1080/15362426.2012.657052.

Bibliography

A Public Trust at Risk: The Heritage Health Index Report on the State of America's Collections.

Washington, D. C.: Heritage Preservation, 2005.

<http://www.heritagepreservation.org/hhi/HHIfull.pdf>.

Babbie, Earl R. *The Basics of Social Research*. 5th ed. Belmont, CA: Cengage Learning, 2011.

Baker, Sara M., and Petty Richard E. "ajority and Minority Influence: Source-position Imbalance as a Determinant of Message Scrutiny." *Journal of Personality and Social Psychology* 67, no. 1 (1994): 5–19.

Bantin, James, and Leah Agne. "Digitizing for Value: A User-Based Strategy for University Archives." *Journal of Archival Organization* 8, no. 3/4 (July 2010): 244–50.

doi:10.1080/15332748.2010.550791.

Bültmann, Barbara, Rachel Hardy, Adrienne Muir, and Clara Victor. "Digitized Content in the UK Research Library and Archives Sector." *Journal of Librarianship and Information Science* 38, no. 2 (January 2006): 105–22.

Burns, Jasmine Elizabeth. "Digital Facsimiles and the Modern Viewer: Medieval Manuscripts and Archival Practice in the Age of New Media." MA thesis, State University of New York at Binghamton, 2013.

<http://search.proquest.com.ezproxy.lib.uwm.edu/pqdtft/docview/1431906037/abstract/CDF267A2DCA43A9PQ/1?accountid=15078>.

- Chen, Ching-chih, Howard D. Wactlar, James Z. Wang, and Kevin Kiernan. "Digital Imagery for Significant Cultural and Historical Materials." *International Journal on Digital Libraries* 5, no. 4 (August 2005): 275–86. doi:10.1007/s00799-004-0097-5.
- Chowdhury, Gobinda. "From Digital Libraries to Digital Preservation Research: The Importance of Users and Context." *Journal of Documentation* 66, no. 2 (2010): 207–23. doi:http://dx.doi.org/10.1108/00220411011023625.
- Confino, Alon. "Collective Memory and Cultural History: Problems of Method." *American Historical Review* 102, no. 5 (December 1997): 1386–1403.
- Connaway, Lynn Silipigni, and Ronald R. Powell. *Basic Research Methods for Librarians*. 5th ed. Santa Barbara: ABC-CLIO, 2010.
- Conway, Paul. "Modes of Seeing: Digitized Photographic Archives and the Experienced User." *American Archivist* 73, no. 2 (Fall/Winter 2010): 425–62.
- Conway, Paul. "Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas." *Library Quarterly* 80, no. 1 (January 2010): 61–79.
- Craig, Barbara L. "Selected Themes in the Literature on Memory and Their Pertinence to Archives." *American Archivist* 65, no. 2 (Fall/Winter 2002): 276–289.
- Crane, Susan. "Memory, Distortion, and History in the Museum." *History & Theory* 36, no. 4 (1997): 44-63.
- Creswell, John W. *Qualitative Inquiry Research Design: Choosing among Five Approaches*. Sage Publications, 2007.

- Davis, Ryan A. "Caught in Franco's Web: Virtual Memories of Prisons During the Franco Regime." *Journal of Spanish Cultural Studies* 9, no. 1 (2008): 89–101.
doi:10.1080/14636200701868035.
- DeRidder, Jody L., and Kathryn G. Matheny. "What Do Researchers Need? Feedback On Use of Online Primary Source Materials." *D-Lib Magazine* 20, no. 7/8 (July 2014).
doi:10.1045/july2014-deridder.
- Dickson, Maggie. "Due Diligence, Futile Effort: Copyright and the Digitization of the Thomas E. Watson Papers." *American Archivist* 73, no. 2 (2010): 626-636. Accessed November 12, 2015. <http://archivists.metapress.com/content/16RH811120280434>.
- Frisch, Michael. "American History and the Structures of Collective Memory: A Modest Exercise in Empirical Iconography." *The Journal of American History* 75, no. 4 (March 1989): 1130–55. doi:10.2307/1908633.
- Gauthereau-Bryson, Lorena. "Global Linking Through Archival Digitization: The Our Americas Archive Partnership and Spanish, Portuguese, and American Women (1876)." *Microform & Digitization Review* 40, no. 1 (March 2011): 25–29. doi:10.1515/mdr.2011.003.
- Gilliland-Swetland, Luke. "The Provenance of a Profession: The Permanence of the Public Archives and Historical Manuscripts Traditions in American Archival History." *American Archivist* 54, no. 2 (April 1, 1991): 160–75.
- Gracy, Karen F. "Distribution and Consumption Patterns of Archival Moving Images in Online Environments." *American Archivist* 75, no. 2 (Fall/Winter 2012): 422–55.
- Hirtle, Peter B. "The Impact of Digitization on Special Collections in Libraries." *Libraries & Culture* 37, no. 1 (Winter 2002): 42–52.

- Hobsbawm, E. J., and T. O. Ranger, eds. *The Invention of Tradition*. Cambridge: Cambridge University Press, 1995.
- Hooper, Lisa, and Donald C. Force. *Keeping Time: An Introduction to Archival Best Practices for Music Librarians*. Middleton, Wis: A-R Editions, Inc, 2014.
- James, Daniel. "Meatpackers, Peronists, and Collective Memory: A View from the South." *American Historical Review* 102, no. 5 (December 1997): 1404–12.
- Johnston, Leslie. "Making It Usable: Developing Personal Collection Tools for Digital Collections." In *I, Digital: Personal Collections in the Digital Era*, edited by Christopher A. Lee, 257–79. Society of American Archivists, 2011.
- Joseph, Pauline, Michael Gregg, and Sally May. "Digitisation of the WA Welcome Wall Collection: A Case Study." *Archives & Manuscripts* 41, no. 3 (November 2013): 183–202. doi:10.1080/01576895.2013.829752.
- Kempner, Joanna. "The Chilling Effect: How Do Researchers React to Controversy?" *PLoS Medicine* 5, no. 11, e222 (November 2008): 1571–8.
- Klein, Kerwin Lee. "On the Emergence of 'Memory' in Historical Discourse." *Representations*, no. 69 (January 2000): 127–50.
- Koltun, Lilly. "The Promise and Threat of Digital Options in an Archival Age." *Archivaria*, no. 47 (April 15, 1999): 114–35.
- Linenthal, Edward Tabor, and Tom Engelhardt, eds. *History Wars : The Enola Gay and Other Battles for the American Past*. 1st ed. New York: Metropolitan Books, 1996.

- Matusiak, Krystyna K., and Myagmar Munkhmandakh. "A Newspaper/Periodical Digitization Project in Mongolia: Creating a Digital Archive of Rare Mongolian Publications." *The Serials Librarian* 57, no. 1–2 (2009): 118–27. doi:10.1080/03615260802669136.
- Mayer, Pamela. "Like a Box of Chocolates: A Case Study of User-Contributed Content at Footnote." *American Archivist* 76, no. 1 (Spring/Summer 2013): 19–46.
- Michel, Peter. "Digitizing Special Collections: To Boldly Go Where We've Been before." *Library Hi Tech* 23, no. 3 (2005): 379–95.
- Middleton, Ken. "Collaborative Digitization Programs: A Multifaceted Approach to Sustainability." *Library Hi Tech* 23, no. 2 (2005): 146–50.
- Millar, Laura. *Archives: Principles and Practices*. New York: Neal-Schuman Publishers, 2010.
- Monks-Leeson, Emily. "Archives on the Internet: Representing Contexts and Provenance from Repository to Website." *American Archivist* 74, no. 1 (Spring/Summer 2011): 38–57.
- Nash, Gary B., Charlotte A. Crabtree, and Ross E. Dunn. *History on Trial: Culture Wars and the Teaching of the Past*. 1st ed. New York: Alfred A. Knopf, 1997.
- National Archives and Records Administration. "Strategy for Digitizing Archival Materials for Public Access, 2015-2024," *National Archives and Records Administration*. 2014. Accessed November 12, 2015. <http://www.archives.gov/digitization/strategy.html>.
- Nora, Pierre. "Between Memory and History: *Les Lieux de Mémoire*." *Representations*, no. 26 (April 1989): 7–24. doi:10.2307/2928520.
- Perry, Claudia A. "Education for Digitization: How Do We Prepare?" *Journal of Academic Librarianship* 31, no. 6 (November 2005): 523–32.

- Pevar, Susan. "Success as a Lone Arranger: Setting Priorities and Getting the Job Done." *Journal of Archival Organization* 3, no. 1 (January 2005): 51–60. doi:10.1300/J201v03n105.
- Preservation committee. "Digitization and Archives." Canadian Council of Archives, 2002.
http://www.cdncouncilarchives.ca/digitization_en.pdf.
- Puglia, Steven, Jeffrey Reed, and Erin Rhodes, *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images*, National Archives and Records Administration, June 2004.
<http://www.archives.gov/preservation/technical/guidelines.pdf>.
- Purday, Jonathan. "Intellectual Property Issues and Europeana, Europe's Digital Library, Museum and Archive." *Legal Information Management* 10, no. 03 (2010): 174–80.
doi:10.1017/S1472669610000678.
- Robb, Evan. "Gleaning Local History: Community-Based Digitization Experiences in Rural Washington." *Microform & Imaging Review* 39, no. 1 (January 2010): 12–7.
doi:10.1515/mfir.2010.002.
- Seadle, Michael. "Selection for Digital Preservation." *Library Hi Tech* 22, no. 2 (2004): 119–21.
- Sider, Gerald M. "Cleansing History: Lawrence, Massachusetts, the Strike for Four Loaves of Bread and No Roses, and the Anthropology of Working-class Consciousness." *Radical History Review* 65 (1996): 48-83. Accessed June 30, 2014. doi:10.1215/01636545-1996-65-48.
- Simon, James T. "World Newspaper Archive." *Slavic & East European Information Resources* 10, no. 1 (January 2009): 82–7. doi:10.1080/15228880902756443.

- Sinn, Donghee. "Impact of Digital Archival Collections on Historical Research." *Journal of the American Society for Information Science and Technology* 63, no. 8 (August 2012): 1521–37. doi:10.1002/asi.22650.
- Sitts, Maxine K., ed. *Handbook for Digital Projects: A Management Tool for Preservation and Access*. 1st ed. Andover, Mass: Northeast Document Conservation Center, 2000.
- Solberg, Janine. "Googling the Archive: Digital Tools and the Practice of History." *Advances in the History of Rhetoric* 15, no. 1 (January 2012): 53–76.
doi:10.1080/15362426.2012.657052.
- Stauffer, Andrew. "The Nineteenth-Century Archive in the Digital Age." *European Romantic Review* 23, no. 3 (June 2012): 335–41.
- United States, Federal Communications Commission, Eighth Broadband Progress Report, (Washington D. C.), Federal Communications Commission, 2012. Accessed July 6, 2014.
https://apps.fcc.gov/edocs_public/attachmatch/FCC-12-90A1.pdf.
- Watson, Andrea, and P. Toby Graham. "CSS Alabama 'Digital Collection': A Special Collections Digitization Project." *The American Archivist* 61, no. 1 (April 1998): 124–34.

APPENDIX A:

Case Study Inquiry Email

Dear **(INSERT CONTACT NAME)**,

Hello. My name is Randy Smith and I am a graduate student in the School of Information Studies at the University of Wisconsin – Milwaukee and am working on a thesis project to determine what personal choices factor into the selection of items a repository decides to digitize for online access. The criteria of this project is to mainly focus on items from a collection that did not have a digital surrogate (a digital facsimile of a physical item meant to be accessed via a computer) created when the rest of a collection was digitized. I am also interested in gaining a better understanding of what choices are made when choosing to digitize one collection over another.

The impetus for this study is to assess what decisions are made when determining which items from a collection an archivist transforms into digital surrogates and whether to make them available on the World Wide Web. The objective of this study is to understand the reasons personal, technical, and institutional factors affect the digitization and publication of archival materials.

According to your archives' website, you have several digital collections available for public access. Are you aware of a situation as described above wherein a personal, technical or institution factor interfered with the creation of a digital surrogate of a specific item(s) from a

collection that was otherwise digitized? If you are aware of such a situation, would you be willing to meet with me, either in person, via Skype, or a phone conversation, to discuss the decisions and processes that were involved in digitizing the materials from the collection(s) and the placing of the digital surrogates online? If yes, please reply to this email and I will arrange a time for the meeting. I realize the holidays can be a hectic time so if there is a date range or specific times that are most preferable for you to meet with me please let me know and I will try to accommodate your request. Once a meeting has been arranged, I will send you a copy of the interview questions I intend to ask.

All information collected during this project will remain confidential unless you provide permission for me to do otherwise.

Thank you for your time and consideration. Please do not hesitate to contact me if you have any questions about this message or my research project.

Sincerely,

Randy Smith

APPENDIX B:

Interview Questions

- 1) What criteria are used to determine which items within your physical collection are to be digitized?
 - a. Do you ever deviate from these criteria?
 - b. What criteria are used to determine which items within your physical collection are to be digitized?
- 2) Do you have examples of instances of collections digitized that did not fit into these criteria?
- 3) Are these criteria always applied during the digitization decision-making process?
- 4) Are there any other factors that you can think that could alter the decision-making process?
- 5) Are all digital surrogates placed on the Internet? Why or why not?
- 6) Are all digital surrogates publically accessible? Why or why not?
- 7) Was the methodology described in question 1 applied to the **(INSERT COLLECTION TITLE)** collection? Why or why not?
- 8) Can you think instances where you might deviate from such methodology?

- 9) You mentioned in the preliminary email that you were aware of instances where an item was deliberately not digitized. Do you **(remember or know why)** such a decision was made?
- 10) Have you digitized items but then decided not to place them online for public access?
- 11) Have you ever digitized an item and placed it online only to later decide to remove it from public internet access? If so, what were the reasons for removing it?
- 12) Is there any chance **(INSERT PREVIOUSLY MENTIONED ITEM)** will be made publicly available via the Internet?
- 13) **(IF THE POLICIES HAVE NOT ALREADY BEEN RETRIEVED)** Do you ever acquire collections with the intent to digitize one or more items from it?
- 14) How many different collections have you placed online?
- 15) Is that the same amount of collections digitized?
- a. If not, please explain.
- 16) About how many total number of items have been digitized to date?
- 17) Have you ever removed digitized items from online access?
- a. If so, what were the items?
 - b. What was the reason for removing them?
- 18) What is the platform utilized to make the digital surrogates available to the public?

19) Are there personnel whose sole task is to digitize?

a. How many?

b. Are they staff members? Volunteers? Some other type of position?

20) How is funding retained for the digitization of items?