The Digitization and Accessibility of Documents: A Case Study at the Rochester Public Library

Kirsten Feigel
klf2504@rit.edu

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
The Rochester Institute of Technology
College of Liberal Arts

The Digitization and Accessibility of Documents:
A Case Study at the Rochester Public Library

A thesis submitted
in partial fulfillment of the
Bachelor of Science degree
in Museum Studies

Performing Arts and Visual Culture Department

By

Kirsten Feigel

May 2015
The members of the Committee approve the thesis of Kirsten Feigel submitted on Thursday, May 21, 2015.

Juilee Decker, Ph.D.
Primary Advisor

Jody Sidlauskas
Secondary Advisor

Tina Lent, Ph.D.
Director, Museum Studies Program
Table of Contents:

I. Abstract ..............................................................................................................................................1

II. Literature Review ..........................................................................................................................2
   a. Brief History of Digitization ........................................................................................................2
   b. Reasons for Digitization ..............................................................................................................3
   c. Purposes and Benefits of Digitization .........................................................................................3
   d. Digitization and Preservation .....................................................................................................4

III. Digitization and Access in American Collections ..................................................................5
   a. Smithsonian Institution ................................................................................................................5
   b. The Library Company of Philadelphia .........................................................................................8
   c. Digital Public Library of America ...............................................................................................9
   d. Library of Congress ....................................................................................................................9

IV. Current Practices ........................................................................................................................10
   a. Step-by-step Process ..................................................................................................................10
   b. Range of Equipment ..................................................................................................................12
   c. Access of Information ................................................................................................................13

V. Case Study: Rochester Public Library ....................................................................................14

VI. Interpretation/Analysis ...............................................................................................................18

VII. Conclusion ...................................................................................................................................20

VIII. Appendix ....................................................................................................................................21

IX. Bibliography ..................................................................................................................................28
I. Abstract

Technology has proven to benefit human experience and achievement in many ways and has become a growing aspect of visitor engagement, documentation, and research in museums. Collections professionals in libraries and archives, as well as museums, are committing to processing artifacts, in many cases, to make them digitally accessible online. Such collections of online assets are classified as digital libraries, which become online tools used for research and descriptive search engines to assist individuals beyond the need of the physical interaction at the place of learning. This thesis provides an overview of the four outline digital repositories based in the United States, including the Smithsonian, the Library Company of Philadelphia, the Digital Public Library of America, and the Library of Congress. Further, I provide a case study of one particular library repository: the Rochester Voices collection which is a part of the Rochester Public Library’s local history site. I describe my experience as a volunteer during the summer 2014 and winter 2015 at the Rochester Public Library’s Local History department where I digitized a collection of manuscript documents, the Raymond L. Rohner papers. I provide a step-by-step process of document digitization and offer reflection and analysis of this experience. This project provides an understanding of how digital libraries are created and maintained, demonstrates how these resources are accessible to and useful for online visitors as well as future researchers, and reveals, ultimately, how libraries, archives, and museums can and should take advantage of new technologies.
II. Literature Review

Digitization is defined as the process of converting the content of physical media, including, but not limited to, periodic articles, books, manuscripts, photographs, film, and audio, to a digital format. This practice has brought significant benefit to libraries, museums, and archives internationally, as digitization, combined with increasing capacity of storage and worldwide internet connection to this information, causes rapid changes from the use of the traditional collection management format of using, managing, and accessing knowledge and information related to these collections.¹

A. Brief History of Digitization

The internet is a widely used tool in today’s society, connecting individuals with unlimited access to information not otherwise physically obtained. The Web had its beginnings in the United States’ Department of Defense Advance Research and Projects Area (DARPA) in the late 1960s, and by the late 1980s, its network was split in two: a private, military-only network and a network used to expand access for research, business, and other personal uses. Libraries and museums did not join the latter network until the late 1980s/early 1990s, when the Institute for Museum and Library Services (IMLS) had become the primary benefactor and source of federal funding, pushing these institutions onto the internet for predominantly educational reasons.² It was about this time when libraries began to organize for and generate

digitization projects for their physically less accessible materials, and to make them available for online research.

**B. Reasons for Digitization**

As new technologies become available, libraries must budget their time, financial resources, and staff members intelligently in order to highlight the idea of digitization in a reasonable direction. Paul Conway gives some examples and reasons behind digitizing collections that can be used internationally. One reason that institutions digitize is to protect the originals. Digital copies can be used for ready reference and not only can these specific copies be limitlessly accessible, they can be protected from serious long-time handling by unprofessional and professional handlers alike. A second reason to digitize is that digital items are a higher quality, higher representation of originals. Old manuscripts and documents may be faded and warn from age and storage, but by digitizing these materials, they can be shown again through computer-based editing software that can increase resolution and make it legible for readers. A third reason is that original documents will be transcended. Digital imaging can be used for many other purposes, such as formatting digitized originals to different programs, adding searchable text, and create an overall sense of agelessness.

**C. Purpose and Benefits of Digitization**

Peter Hirtle describes the benefits of digitizing a special collection as a backbone for new types of research, accessibility, and creating new users. “Digitization has brought significant benefits to the users of special collections. It will also challenge the relative value given to paper

---


In many cases, digitizing can preserve fragile materials by reducing the handling by visitors and other librarians, as well as improve management of all collections, whether by using databases and or other online resources. In a digital world, preservation is the creation of digital products worth maintaining over time. Digitization coupled with preservation, affords libraries, archives, and museums the opportunity to place their original material in appropriate archival and secure storage facilities to which access can be restricted given the existence of the digital surrogate. Prior to World War II, preservation was simply defined as collecting and limiting the uses of primary documents in fear of damage, theft, and misuse by handlers through the years. Modern preservation techniques claim and confirm that preservation and access are mutually reinforcing ideas. Preservation is widely practiced today so that access may be possible, and collections would be digitally protected with the use of external hard drives, backup computer systems, and extensive online maintenance. Emulation is a process that involves taking technological steps in order to enable future computer systems to read digital information with minimal loss of original look, feel, and functionality of collections. Migration involves the transfer of data from one technological environment to another, updated version, but migration has its downfalls. There are possibilities that with each migration, data could be lost because of transfer complications from one format to another. Another problem with migration that libraries, archives, and museums can face is that as

---

digital collections increase in size, this practice can become very time consuming and expensive. But today, migration is the only practical strategy in order to preserve large scale digital archives over time.

Members of the library, museum, and archive community must take into consideration the capabilities certain technological advances have in order to scan documents and prevent them from becoming obsolete through the use of preservation. They must understand that all conversion and digitization tools have limitations on size of source documents they can handle at whatever resolution. These digitization tools are not inexpensive; financial budgeting is increasingly important as new technology is brought into the workspace, forcing other machines and digital imaging software to become unusable. Overall, digital preservation should concern any library, museum, or archive that is considering a digitization project of the limitations, costs, and timescale, even if improving individual access is the primary motivating factor.⁹

III. Digitization and Access in American Collections

Digitization practices started in the late 1990s worldwide, creating online access to anyone with an internet connection. Native to America are four institutions, both physical and online, whose missions are to create access to their collections in any means possible to entertain, educate, and involve the public.

A. The Smithsonian Institution

The Smithsonian Institution is widely recognized as the world’s largest museum and research complex, housing 137 million artifacts, hosting nineteen museums and galleries, nine research facilities, and the National Zoological Park. This institution was founded in 1846 in the

metro area of Washington D.C. and has been ever since a place of research and knowledge, free of charge. G. Wayne Clough, the secretary of this institution from 2008 to 2014, issued a digital publication titled “Best of Both Worlds: Museums, Libraries, and Archives in a Digital Age”, in which he explains the institution’s history of digital practices and future outlooks. The National Museum of Natural History began to digitize their botanical specimens as early as the 1970s, and not long after, the Smithsonian American Art Museum followed suit, developing the first searchable database for American art. As technology advanced, the officials of the institution had to reinforce their mission: “the increase and diffusion of knowledge”, to a better and higher understanding for those entering the technological world and who were still physically entering the strip of museums to learn.

In 1995, the Smithsonian processed a strategic plan, one that would start their journey of large-scale digitization efforts in order to reach a broader audience. This led to the creation of the “Collections Information Systems Pool”, a committee which helped raise and provide funds for the equipment and digitization practices they would soon use. In 2008, the year Clough started at the museum, digital technology was ubiquitous; computers were becoming smaller and more interactive, mobile devices were evolving, and younger individuals were starting their own digital journeys. The following year, another strategic plan was created for the Smithsonian, one that would surely lay down the basis of what the institution planned on doing in response to changing technologies and how other learning institutions around the world were taking advantage of them. In order to reach the audience on a broader spectrum, the Smithsonian 2.0 conference in 2010 explored possibilities of Web-based technologies in their midst, and in result, the Smithsonian seed grant program was established in order to fund innovated applications for digital technology.
The pioneers of the Smithsonian digital journey, the Botany department at the National Museum of Natural History and the Smithsonian American Art Museum, both have flourished with these new technologies since the 1970s. As of today, the National Museum of Natural History has digitized more than 1 million of an estimated 5 million botanical specimens. The Smithsonian American Art Museum holds digital records for its entire art collection and images for more than eighty percent can be found on their website. The Center for Folklife and Cultural Heritage research institution of the Smithsonian has digitized sixteen percent of their audio recordings that are available for purchase on iTunes. Working with the Gale Cengage publishing company, the Smithsonian has been able to digitize their *Smithsonian* and *Air & Space* magazines, which, among many others, can be found via collections search center on the institution’s main website.

Clough describes the task of digitizing to be a “daunting” one, which requires years of funding, equipment, and institutional planning. When it comes to their 137 million artifacts, the institution must make sacrifices and follow priority protocol, in which the institution developed a comprehensive survey for the officials. This survey resulted in the agreement that out of number, that about 13 million, or 10%, should be considered as priority for digitization.

The Smithsonian Institution has given access to its patrons ever since it was founded, and focuses primarily on educating the public. With its free admission to 19 museums and galleries, nine research facilities, and the National Zoological Park, this prized center of knowledge has been, and will continue to focus their mission primarily on educating the public in a more informal experience. As digitization came into play for museums, libraries, and archives worldwide, the Smithsonian was not an institution to sit back and judge from afar. Their main
website and digital collections have been seen by millions of people during the years, those who have been publically engaged since the day the doors opened.\textsuperscript{10}

\textbf{B. The Library Company of Philadelphia (LCP)}

Not as widely known as the Smithsonian Institution, the Library Company of Philadelphia is America’s first successful lending library and oldest cultural institution. Founded in 1731 by Benjamin Franklin, this educational facility started out as an inexpensive subscription library, for books were not widely available to all at the time. Franklin would convince members of his inner circle to pool their resources in order to purchase a collection of books from overseas and allow access to the public. In the 1850s, the LCP became the largest public library in America. Today, the LCP is an independent research library that concentrates on American society and culture from the seventeenth to nineteenth centuries.\textsuperscript{11}

Patrons are welcome to use both the physical library space and the online library for accessibility. The LCP’s online catalog, WolfPAC, allows researchers to search nearly 100\% of LCP’s rare book collection, as well as the collections of neighboring institutions in Philadelphia.\textsuperscript{12} ImPAC, LCP’s digital collections catalog, currently contains 35 digital collections including almost 5000 images, all available for download and viewing by online visitors.\textsuperscript{13}

\textsuperscript{11} “About” The Library Company of Philadelphia, \texttt{http://www.librarycompany.org/about/index.htm}
\textsuperscript{12} These neighboring institutions are the Athenaeum of Philadelphia, The Philadelphia Horticultural Society, The Philadelphia Museum of Art, The Rosenbach Library and Museum, and The Wagner Free Institute of Science, but these are beyond the scope of this paper.
\textsuperscript{13} “Overview.” The Library Company of Philadelphia, \texttt{http://www.librarycompany.org/collections/index.htm}.\hfill 8
C. Digital Public Library of America (DPLA)

The Digital Public Library of America is a registered non-profit organization and registered as a library in the state of Massachusetts. It is backed by a range of organizations, including the Library of Congress, HathiTrust, and Internet Archive, each of which had provided books, images, historic records, and audiovisual materials to anyone with an internet connection.

In October of 2010, the planning process for DPLA started at a conference in Cambridge, Massachusetts, where over 40 leaders from libraries, foundations, academia, and technology projects agreed to work together to create “an open, distributed network of comprehensive online resources that would draw on the nation’s living heritage from libraries, universities, archives, and museums in order to educate, inform, and empower everyone in current and future generations.” With their search engine, Portal, and application programming interface, Platform, DPLA can easily bring the public together with American’s libraries, archives, and museums, to make them freely accessible and available to the world.14

D. Library of Congress (LOC)

“The Library’s mission is to support the Congress in fulfilling its constitutional duties and to further the progress of knowledge and creativity for the benefit of the American people.” As the nation’s oldest federal cultural institution and largest library in the world, the LOC is home to millions of books, recordings, photographs, maps and manuscripts in its collections.15

The LOC was established through an act of Congress in the year 1800 when President John Adams signed a bill that transferred the seat of the government from Philadelphia to

---

15 “About the Library.” About the Library. Library of Congress http://www.loc.gov/about/.
Washington DC. The legislation created this reference library solely for the use of Congress, housing “such books as may be necessary for the use of Congress—and for putting up a suitable apartment for containing them therein…” Five-thousand dollars later, the original library was housed in the new Capitol building until August of 1814, when it was set ablaze by invading British troops. Thomas Jefferson, as former president, offered his personal library as replacement, having collected 6,487 books in the time frame of fifty years.¹⁶

Today, the LOC has a collection of more than 158 million items: more than 36 million catalogued books and other print materials in 460 languages, more than 69 million manuscripts, is home to the largest rare book collection in North America, and is the world’s largest collection of legal materials, films, maps, sound recordings, and maps. The Library has also made digitized versions of their collection materials available online, starting their digitization project in 1994, concentrating on its most rare collections and those unavailable anywhere else. The digitized collections contain photographs, manuscripts, maps, sound recordings, motion pictures, and books, as well as Web sites. The LOC maintains and promotes the use of the digital library standards and provides to the public research and reference services daily.¹⁷

IV. Current Practices

Libraries, archives, and museums everywhere are taking the step forward into digitizing their collections to help maximize the meaning of each institution’s mission and vision. As some of these institutions practice digitization in different ways, such as using book scanners or taking digital photographs by hand, the universal practice of digitization promises the increase of online accessibility for generations.

---

A. Step-by-Step Process

Digitizing a library, archive, or museum collection for online use and accessibility is nothing more than a project that requires individuals who can be flexible and trusted to do assigned tasks. Like every other project, digitization requires a time frame, funding, and dedicated staff members who have a clear understanding of an objective and a result in mind.

For larger, more well-to-do libraries, digital imaging projects may have advantages backing them up. These institutions may have on staff dedicated programmers, luxuries of costly commercial software, and multiple librarians with various backgrounds and specializations. However, lessons must be taught, rules must be followed, and ideas must be put into consideration before these projects can be started. For one, librarians and other officials must accept their limitations; being realistic in completing their goals in a timely manner is important in order to save financial resources and sanity, and no one person can do it alone. A second lesson is that not all software and hardware is really free. Content management systems and content delivery systems often have financial strings attached, and libraries must set aside resources in order to acquire such materials, such as Umbraco and Google Analytics. Another tip to follow is to take advantage of fundraising and grant writing opportunities. Budget cuts may become an unseen issue, especially when money is tight.

Libraries need to be sure to start small and simply. For digitizing certain materials, whether tasks include photographing for library websites, or scanning documents and other scholarly materials, some mistakes may not be easily corrected. By starting with smaller, clustered collections, the project will become more manageable and provide a sense of confidence of work ethic for future projects. A fifth lesson for digitizing any library collection is
learning to juggle with other projects, interrelated or not, in order to maintain balance. While handling multiple projects at once may seem like a stressful and less organized path to take, it does have its benefits. During the decision making process that evolves around one certain collection, ideas and processes may be allocated for future projects and influence work flow. This way, when hurdles are met with one project, librarians can identify them, document them, and learn from them so they do not run into the same problem a second time.

Collaboration is highly regarded when it comes to these kinds of projects. Whether it is through the same department of the same library, or a different institution altogether, bringing together a team of individuals who have the same outlook and perspective on the projects final outcome will become beneficial for the whole team. One last tip of digitizing collections is another self-realization and realistic one; the progress of the project will be slower than one might expect. Whenever there is an anticipated time frame assigned with a project, librarians should double it, for digitization is a process where it pays to get it right the first time.18

B. Range of Equipment

Depending on the digitization project at hand, whether it is a scanning project or one that requires photography, there is a range of equipment that all museums, libraries, and archives must have on hand in order to complete these digital initiatives.

A well-equipped digitization station will have a range of computer-based components and software such as a Windows XP Professional or a Mac. A 2.13-GHz Intel Core 2 Duo Processor, or one much faster, is a reliable processor, as projects will be completed in a timely manner and without a lagging hard drive. Any such book scanner, whether it is a Plustek

---

OpticBook 3600 Book Scanner, or an Indus Color Book Scanner, will be beneficial when it comes to rapidly scanning pages of large books. An Epson Expression Flatbed Scanner, or any other flatbed scanner, is used for digitizing flat materials, such as letters, diary entries, and other free manuscripts. Adobe Photoshop and other photo manipulation software is used to edit problem scans and even format them as different files in order for the computer to read certain types while on the internet.\footnote{Regina L. Avila, Susan Sanders, and Keith Martin, "Tips and Tools for Digitizing a Museum Collection," \textit{Online}, November/December 2011.}

C. Access of Information

With the Age of Information currently taking place in society today, access to websites, databases, and articles is an easy task for any individual with even the slightest knowledge of technology and an internet connection. Today, the massive information choice and continuous twenty-four-seven connectivity to almost anywhere on the Web is a fascinating privilege to scholars and those uneducated alike. Daniel Nicholas concludes that every individual is a scholar in their own way, now that scholarly information is so freely available via educational websites and digital libraries. Students are logging onto accounts that their university libraries are subscribed to, and in the comfort of their own home, anybody can learn and educate themselves formally or informally. Although this revolution of digital libraries and other online educational components have been around for a decade, growth rates on certain publishing databases have exceeded fifty percent, with acknowledgements to improved wireless connection, increasing students worldwide, increased digitization efforts from libraries and archives, and the preference of the young having everything new, digital, and technologically advanced.\footnote{David Nicholas, "The Virtual Scholar: The Hard and Evident Truth," in \textit{Digital Library Futures: User Perspectives and Institutional Strategies} (International Federation of Library Associations and Institutions, 2010).}
According to Daniel Teruggi, there are four types of scholars that interact and use online materials for research and other means. The first type of user is the general user; a patron that use a system in order to seek out some kind of result, just for curiosity, for educational need, or for entertainment. The second type of user is the specialized user, or the individual that become highly acquainted with a specific online tool or website without being a complete professional. The professional user is the scholar that uses these tools and websites and relies on them heavily in order to obtain specific results. The final user is the representative user, a person with extensive knowledge on a subject, and is trusted to define requirements for other users who may have less knowledge on the subject.  

V. Personal Case Study: The Rochester Public Library

The Local History and Genealogy Department at the Rochester Public Library houses well over five-million documents and photographs documenting Rochester and the Genesee Valley’s history over the past two-hundred years. Not only can generations of family members sift through old newspaper clippings and housing deeds regarding ancestors, but volunteers can handle and research from these Rochester collections under the watchful looks of the trained library staff. During the summer of 2014, I worked under the Deputy City Historian, Michelle Finn, to digitize a collection for the library-run website, Rochester Voices. This website currently hosts a limited amount of collections, ranging from documents written by famous Rochesterians to oral histories from decades past. It targets grades kindergarten through twelve; an educational experience for teacher, student, and researcher.

---

See Figure 5 in appendix.
As part of my volunteer experience, I was charged with working on the Raymond L. Rohner papers, an estimated total of 1400 handwritten letters dated in two chunks: from 1936 to 1938, and from 1943 to 1945. Rohner was a medic stationed at Camp Pickett, Virginia during World War II. The letters serve as a first-hand account of wartime conditions and military procedures. In the earlier letters, dated from 1936 to 1938, a courtship between Rohner and his soon-to-be wife, Else, is painted clearly for the readers. Rohner was known for writing in a humorous and detailed style, adding small drawings to illustrate the narrative. After the war, Rohner and his wife picked up and left New York City for the Rochester area, where he worked as the Head Brew Master at Genesee Brewery for a number of years.23

In terms of the setting for my project, there was one desk at which to work directly outside of the office of the City Historian, Christine Ridarsky. At the time of my arrival at the Rochester Public Library, time became a virtue for both me and the interns working for college credit. Our technological set up was basic. We had a computer, an HP desktop PC and a desktop scanner. The Epson V700 Photo/V750 Pro24 came with its own Epson scan software which allowed for ranging dpi, preview before scan, and access to Media Impression Photo Viewer for Kodak photo editing. As I played around with simple scraps of paper to get used to the scanner, I became familiar with dpi, or dots per inch, quite quickly. Depending on dpi size, the larger the dpi, the higher quality scans, and longer scans time. The Rochester Public Library has certain specifications for their online documents, and they are as follows: with TIFFs, a 400-dpi was recommended. For JPGs, full size and thumbnails, they were scanned at a 72-dpi. I was scheduled to scan TIFFs, JPGs at full size, JPGs at thumbnail size, and PDFs.

23 *The Raymond L. Rohner Papers, Local History and Genealogy Division, Central Library of Rochester and Monroe County, NY.*
24 See Figure 1 in appendix.
Beyond the area where I volunteered, the library has a central digitization department, also located on the second floor. The Digitization Department has more sophisticated equipment to process materials from the collections. To familiarize myself with this area, I shadowed a digital librarian who handles the larger equipment in the equally larger room. She led me to the Indus Color Book Scanner 5005\textsuperscript{25}, which seemed, in my opinion, easier to tackle than the desktop scanner, despite its more extensive software. The Indus Color Book Scanner 5005 was known to scan with a precision like no other, of course, if used correctly. It is a large format face-up color and grayscale scanner for the larger collections, such as books and flat documents. Because of the dual LED lighting programmed into the scanner, one must close any blinds to not allow for extra light to enter the room, for if not done correctly, a pink hue will be left on the documents. The motorized book cradle on the scanner bed has a glass plate to flatten books and documents for perfect results. These cradle halves are also independent and can be raised and lowered if deemed necessary. It is also possible to change the lighting and focus of the scanner right on the machine instead of having to wing it with a simple desktop scanner. The librarian then showed me the Kirtas Book Scanner 1200\textsuperscript{26}, a machine the library is most vocal about having in their collection. While the Indus resembles an overhead projector, the Kirtas comes with mirrors and vacuum-tight sealants that hold each separate page down for easier digitization. There are many “how-to” videos online regarding this specific machine that shows the process the machine makes from within it.\textsuperscript{27}

The Indus scanner appeared to be a much better device than the desktop scanner I was working with, for the Epson was a hit-or-miss scanner. The Epson and its online component

\textsuperscript{25} See Figure 2 in appendix.
\textsuperscript{26} See Figure 3 in appendix.
\textsuperscript{27} Kirtas Book Scanner: \url{http://kirtas.com/bookscanners.php}.
software would occasionally crash in the middle of scanning which interrupted my workflow and impeded progress. Moreover, as I am not a very particularly patient person when it comes to faulty technologies and computers, I found myself frustrated because I had to reboot the system entirely after crashes caused by attempting to use the scanner. This meant that my time was spent inefficiently: where I might have been able to process three folders worth of letters with proper equipment, I was only able to process one. A second factor adding to my frustration was my schedule, a fault on my part. I was only able to work at the library once a week during the summer because I was also interning at the Strong National Museum of Play over the summer 2014.

It became clear to me there was no possible way I would be able to finish scanning the entirety of the box, or even to post the TIFFs online during the summer 2014. Therefore, I returned to volunteer at the library over winter 2014-2015 break to scan additional letters. There were days where I was shaky about remembering the slightest specifics from the summer, but I caught up quickly. By the end of winter break, I completed the scanning of three-quarters of the box. I also formatted the files and posted them to the content management system.

A content management system (CMS) is a computer application that allows for publishing, editing, and modifying content for online content such as blogs and other websites. CMS are also known as the “backend” of a website. The Rochester Public Library uses Umbraco 4 for their CMS.²⁸ I managed to do a quick walk through with Michelle Finn to get an introduction on how to use the system. To first create a folder for the collection I was working on, I had to create pages for the Rohner papers. Then I uploaded the scans from the library’s

See Figure 4 in appendix.
network and put them under the “media” tab of Umbraco 4. For the PDF files, I did not have to create subfolders for them, but for the TIFFs and both JPGs, I did. This took up a lot of time just having to create subfolders for over 180 scans, but overall, this was all done and over with in a time period of two days.

With my time well spent and the experience I have had, digitization was certainly something that interested me. I have also come to know that there are a few people who work on a certain collection, especially when it comes to creating an online presence for one. As I worked on the digitization and scanning of the Rohner collection, there are also content writers and transcribers who provided context for this collection. They gave a brief background of each letter, and transcribed what is written.

While the Rohner collection will not be on Rochester Voices until later this year, I am gratified by the work that I was able to do towards the digitization of this collection. Moreover, I am able to understand what goes into creating an online collection. This hands-on experience over several months will surely reflect on my future work as I continue my education in the fall 2015.

VI. Interpretation/Analysis

Making an online presence for special collections is just the first step in creating easier accessibility for researchers, teachers, and students. Overall, the Rochester Voices website is live and easy to log onto through the library’s website and through the Monroe County Library System. Rochester Voices is a project that Deputy City Historian Michelle Finn has been working on in the past year, and is an interactive, mobile-friendly website that is designed to fully engage the kindergarten through twelve grade audience as well as the general public. This
website gives visitors the chance to explore stories and timelines regarding famous Rochesterians and their personal, first-hand accounts through handwritten letters and oral histories. Rochester Voices is categorized as an online special collections catalog which holds interpreted digital exhibits and an exploratory learning laboratory. The funding that supported the creation and current works of Rochester Voices came from the Gleason Family Foundation. Along with this funding, some partner organizations that contributed to the making of this site include the Rochester Museum and Science Center, and the East Rochester Historian’s Office.29

In order to analyze the success of accessibility for the website, Rochester Voices uses the data collecting platform, Google Analytics. Google Analytics is a web service that provides statistics and basic tools for search engine optimization, and features different website performance levels according to personal company strategy. Finn uses this platform to track the use and her website’s access in order to interpret location, page returns, time each user allotted for each different page, and through what mobile and technological devices individuals are using to log onto this website. With my experience working with Rochester Voices and Finn, I collected this data from the Google Analytics to create a clearer understanding of current accessibility of the website.

Graph 1, located in the appendix, is the amount of sessions, or the amount of times this website was logged onto, from a variety of different countries around the world. Graph 2, titled “Amount of Sessions With Technological Devices”, lists twenty-six different cellular phones, tablets, and other handheld devices that have access to the internet, and the amount of times Rochester Voices was accessed by these devices. Graph 3, also including the same technological devices, supports and analyzes the amount of pages the individual visited during their one session.

session. Lastly, Graph 4 shows how long, on average, each session took while using these handheld technologies. The overall analysis of these statistics is sadly underdeveloped, but in the sake of Rochester Voices, the website has only been live one year this past March. As more content is put onto the website, and more funding is provided, it is to say, without a doubt, that these statistics will increase during the upcoming years.

VII. Conclusion

Libraries everywhere are starting to digitize materials for online access as the technology advances through the years. This is a growing occupation and will continue to grow as more tech-savvy users put down the books and log onto library websites which hold first-hand accounts of certain individuals. The Rochester Public Library and the libraries and museums mentioned in this thesis demonstrate how museums and libraries are utilizing digitization practices for greater access and other purposes.

Technological advances are occurring in our everyday society, with new functionalities and hard drives, some members of the community may not be able to catch up. But in order to increase access to online materials through institution websites such as libraries, museums, and archives, the preservation of the physical object to the digital one is not a lack in judgement but an increasing success in the world today.
VIII. Appendix

Figure 1: Epson Perfection V700 Photo Scanner

Photo credit: Kirsten Feigel, summer 2014
Figure 2: Indus Color Book Scanner 5005
Photo credit: Missouri University Library

Figure 3: Kirtas Book Scanner 1200
Photo credit: Smithsonian Library and YouTube
Figure 4: Umbraco 4 screen capture

Photo credit: CMS Critic

Figure 5: Rochester Voices Home Page
Graph 1: Amount of Sessions per Country
Data collected from Google Analytics
Graph 2: Amount of Sessions with Technological Devices

Amount of Sessions with Technological Devices

- ZTE Momodesign MD Droid
- Samsung SM-G900V Galaxy S5
- Samsung SM-G850A Galaxy Alpha
- Samsung SCH-R970 Galaxy S IV
- Samsung SCH-I605 Galaxy Note II
- Samsung SCH-I337 Galaxy S IV
- Samsung GT-P3113 Galaxy Tab 2 7.0
- Samsung GT-I9505 Galaxy S IV
- Nokia X2-02
- Motorola XT 1030 Droid Mini
- LG VS985
- LG LW770 Optimus Regard
- LG LS840 Viper 4G LTE
- HTC ADR6350 Droid Incredible 2
- Samsung SM-P609 Galaxy Note 10.1 2014
- Samsung SCH-I535 Galaxy S III
- Apple iPod
- Samsung SPH-L900 Galaxy Note II
- Samsung SM-N900V Galaxy Note 3
- Motorola XT1060 Moto X
- Google Nexus 5
- (not set)
- LG VS880 Vista
- Motorola XT907 DROID RAZR M 4G LTE
- Apple iPad
- Apple iPhone
Graph 3: Pages per Session via Mobile Devices

Pages per Session

- ZTE Momodesign MD Droid
- Samsung SM-G900V Galaxy S5
- Samsung SM-G850A Galaxy Alpha
- Samsung SCH-R970 Galaxy S IV
- Samsung SCH-I605 Galaxy Note II
- Samsung SCH-I337 Galaxy S IV
- Samsung GT-P3113 Galaxy Tab 7.0
- Samsung GT-I9505 Galaxy S IV
- Nokia X2-02
- Motorola XT 1030 Droid Mini
- LG V5985
- LG LW770 Optimus Regard
- LG LS840 Viper 4G LTE
- HTC ADR6350 Droid Incredible 2
- Samsung SM-P609 Galaxy Note 10.1 2014
- Samsung SCH-I535 Galaxy S III
- Apple iPod
- Samsung SPH-L900 Galaxy Note II
- Samsung SM-N900V Galaxy Note 3
- Motorola XT1060 Moto X
- Google Nexus 5
- (not set)
- LG VS880 Vista
- Motorola XT907 DROID RAZR M 4G LTE
- Apple iPad
- Apple iPhone

Pages/Session
Graph 4: Average Session Duration via Mobile Devices

Average Session Duration

- ZTE Momodesign MD Droid
- Samsung SM-G900V Galaxy S5
- Samsung SM-G850A Galaxy Alpha
- Samsung SCH-R970 Galaxy S IV
- Samsung SCH-I605 Galaxy Note II
- Samsung SCH-I337 Galaxy S IV
- Samsung GT-P3113 Galaxy Tab 2 7.0
- Samsung GT-I9505 Galaxy S IV
- Nokia X2-02
- Motorola XT 1030 Droid Mini
- LG VS985
- LG LW770 Optimus Regard
- LG LS840 Viper 4G LTE
- HTC ADR6350 Droid Incredible 2
- Samsung SM-P609 Galaxy Note 10.1...
- Samsung SCH-I535 Galaxy S III
- Apple iPod
- Samsung SPH-L900 Galaxy Note II
- Samsung SM-N900V Galaxy Note 3
- Motorola XT1060 Moto X
- Google Nexus 5
- (not set)
- LG VS880 Vista
- Motorola XT907 DROID RAZR M 4G LTE
- Apple iPad
- Apple iPhone

Average Session Duration

0:00:00 0:07:12 0:14:24 0:21:36
IX. Bibliography


*The Raymond L. Rohner Papers, Local History and Genealogy Division, Central Library of Rochester and Monroe County, NY.*