

ABSTRACT

“Railroad Iron Is A Magician's Rod’: Preserving And Presenting Historic Railroad Maps Through The ‘Magic’ Of Digitization”

Eric S. Ames

Project Chairperson: Julie L. Holcomb

As digitization projects have become increasingly common, museums of all sizes have fallen under the spell of putting their archival materials online. But the process involves a great deal of advanced planning, sifting through software and hardware options, and finding the scanners and staff to carry them out. Often, small to medium-sized museums lack the resources to carry out a successful digitization project on their own. This project outlines a collaboration between the author and the Railroad and Heritage Museum (RHM) in Temple, Texas. The RHM is home to a collection of late 19th- and early 20th-century railroad engineering survey maps that are often requested by researchers. The RHM had expressed interest in digitizing them and placing them online, but lacked the staff and scanners necessary to do so. Using the resources of the Riley Digitization Center at Baylor University, the author digitized 326 maps and placed them online using Flickr (a digital asset management solution) and Pachyderm (a Flash-based Web site authoring software created for use by museums). This project details the steps involved in that process, as well as some recommendations for museums of modest size that seek to create their own online presence with limited resources.

“Railroad Iron Is A Magician's Rod’: Preserving And Presenting Historic Railroad Maps Through The
‘Magic’ Of Digitization”

A Graduate Project Submitted to the Graduate Faculty of
Baylor University
In Partial Fulfillment of the
Requirements for the Degree
of
Master of Arts

By
Eric S. Ames

Waco, Texas
May 2010

Copyright © 2010 by Eric S. Ames

All rights reserved

TABLE OF CONTENTS

LIST OF FIGURES	iv
ACKNOWLEDGMENTS	v
DEDICATION	vi
CHAPTER ONE	
Digitization Theory and Development	1
CHAPTER TWO	
Selecting a Museum Partner	22
CHAPTER THREE	
The Digitization Process	28
CHAPTER FOUR	
Creating an Online Presence	35
CHAPTER FIVE	
Conclusions	53
APPENDIX A	56
Selections from the RHM Houston Public Library Collection	57
APPENDIX B	
Survey information	
APPENDIX C	
Pachyderm Site Screen Captures	

LIST OF FIGURES

Fig. 1. Granularity in metadata records	9
Fig. 2. Cruse SLT-285 large format scanner	29
Fig. 3. Sample preservation master image and access derivative	33
Fig. 4. The RHM homepage in Flickr	41
Fig. 5. Item record view of map in Flickr	43
Fig. 6. The Flickr map	45
Fig. 7. The online exhibit homepage, via Pachyderm	51
Fig. 8. The “Want to Learn More?” slide from the online exhibit	52

ACKNOWLEDGMENTS

This project simply would not have come to fruition without the support and guidance of many outstanding individuals. I am grateful to Professor Julie Holcomb for her mentorship and unwavering positivity as my project chairperson. My thanks also to Dr. Ellie Caston, Dr. Thomas Charlton and Beth Tice for serving most admirably as my project committee.

Thanks especially to Judy Covington and Craig Ordner at the Railroad and Heritage Museum for partnering with me on this journey. Their professionalism, helpfulness, and willingness to allow 700 maps to travel from Temple to Waco made this whole thing possible.

A big thanks also to all the museum directors who participated in my survey, and to TAM executive director Ruth Ann Rugg for her help getting the survey out to Texas museum professionals.

My thanks also to Tim Logan, Billie Peterson-Lugo and Darryl Stuhr of the Baylor University Electronic Library for their unwavering support and magnanimous flexibility during my three years of graduate study.

Lastly – and by no means not least – my deepest thanks to my wife Amy, my family, and my dear friends the Gracks for their love and support.

Dedicated to

My daughter, the biggest thing to happen for me in 2010.

CHAPTER ONE

Digitization Theory and Development

Perhaps no word in the milieu of museum terminology conjures up more misconceptions, preconceived notions, and shades of gray than “digitization.” Alternately a bogeyman or a savior, an unwieldy behemoth or a magic bullet, digitization exists in a world replete with seemingly endless choices – full color or bitonal? Local storage or off-site? Open access or on-site only? – that often create as many stumbling blocks as they do opportunities. Add in the built-in costs involved with digitization projects, the training and staff time required to carry them out, and the uncertain nature of how end-users will utilize your offerings, and the dream of digitization can quickly turn to a nightmare for a museum with limited resources.

But over the course of the past decade, thanks to the work of dedicated groups of information specialists, some signs of consensus have begun to emerge. Favored file formats have been adopted, strategies for storage embraced, and solutions for online presentation are increasingly available. And for small museums, the promising nature of collaboration with a larger entity may mean enhanced access to unique collections without the burden of investing in new equipment, specialized staff, or long-term storage.

This project describes one such collaboration between the author and the Railroad and Heritage Museum in Temple, Texas. Utilizing the equipment and expertise of the Ray I. Riley Digitization Center at Baylor University, a collection of unique materials was digitized, placed online, and presented in an innovative format for use by researchers around the world. What follows in chapter one is a basic overview of digitization as a strategic solution for museums and libraries, a discussion of challenges and opportunities for museums that choose to digitize, and a look at some best practices. It is not intended

to be an exhaustive overview of the digitization process; rather, it provides the foundation upon which the steps in this project were built.

A brief overview of digitization

The advent of increasingly powerful and more affordable scanning technology in the mid- to late-1990s brought with it an attendant increase in the number of institutions that viewed digitization as a valuable tool in the quest to present information to the public. A similar expansion in access to – and reliance on – the Internet as a means of delivering material to interested parties created fertile ground for libraries, archives, museums and universities to embrace digitization projects, often without much thought beyond the initial “How do we get these items scanned?” phase.

Among the earliest and most powerful rationales behind digitization projects is the ability to preserve, in digital form, the essential substance of a physical (analog) object. This need to preserve the essence of an object as insurance against the loss of the physical is a powerful motivator for institutions charged with protecting our cultural heritage. Digitizing a Civil War-era letter, for example, means that future scholars can access the writer’s words, see his penmanship, even examine minute details like ink stains and minor damage, all through the convenience of a computer monitor located half a world away. In the event of a catastrophic loss to the original – through fire, water damage, theft, even eventual dissolution into dust – the holding institution can provide scholars with access to the digital version, or, if necessary, reprints of the original document.

The ability to provide greater access to an institution’s collections is an equally important aspect of digitized assets. Especially in situations where an object is fragile or easily damaged, digitization gives museums the opportunity to provide virtual access to objects that might otherwise never be available to the public. Before the proliferation of digitized resources available on the Internet, access to digital

images was often limited to local computers accessible only on-site. Other options for delivering images of physical objects to visitors were through museums' use of interactive exhibits or video presentations. Although physically tethered to the host museum and available only to in-person visitors, these methods could be argued to be the forerunners of modern museums' use of digitized objects in online exhibits.

Early challenges

A major stumbling block to early scanning projects was cost. A standard 8.5"x11" flatbed scanner could sell for hundreds to nearly \$1,000 dollars in the early days of desktop scanning; similar models today can be purchased for under \$100. Scanners were seen as a luxury reserved for only the wealthiest institutions, especially compared to the relatively cheap process of converting archival materials to microfilm.

In addition to the attractive cost-to-result ratio offered by microfilm, archiving institutions also cited microfilm's stability, the low cost to store it, and the fact that it can be read with just a lens and a light source as reasons to stick with microfilm versus switching to digitization. But as digitization's long-term viability began to improve, and as prices began to moderate, more and more information professionals began to see advantages to digitization over microfilm, namely its accessibility without requiring a specialized reader, the ability to easily copy it to multiple storage locations, and the fidelity to the original source that comes with digitization. Although some institutions currently utilize a hybrid approach that uses both digitization (for access and other derivatives) and microfilm (for preservation), the trend of switching from microfilm to digital continues to swing toward digitization.¹

¹ Anne R. Kenney, *Digital to Microfilm Conversion: A Demonstration Project 1994-1996* (Ithaca, NY: Cornell University Library, 2001), 3.

Additional costs were required to purchase computers and software necessary for the storage and manipulation of digitized images, a financial burden that many smaller to mid-sized museums could not meet. While many of the high-end, very specialized scanners can still cost more than \$100,000, the boom in affordable, versatile flatbed scanners has leveled the playing field for even small institutions seeking access to the digitization game.

Coupled with cost was the relative newness and untried nature of the technology itself. Early desktop scanners were often large, clunky, slow, and noisy, hardly a combination of attributes hardly likely to sell a suspicious collections manager on the idea of investing hundreds of dollars on the latest technological fad. It wasn't until the late 1990s, when desktop flatbed scanners had greatly increased in effectiveness and the scope of their capabilities, that they were widely adopted by even the most technophobic institutions. Today's scanners are capable of up to 12,000 dpi (dots per inch) scans in full color, a far cry from the 300 dpi maximum scans and very slow completion times of a mere decade ago.

As with all things related to technology, the situation regarding both the cost and efficacy of scanners improved greatly in a relatively short time. The mistrusting mindset adopted by many in the library, archive, and museum world was slower to change, however. A common refrain heard in the early days of digitization was, "If we scan our materials and put them on the Internet, no one will want to visit our archive/museum/library to see the real stuff." While emerging evidence² suggests the opposite is true, it can be difficult to overcome this perception that may be prevalent among the established professionals of the archival and museum professions.

² Jose-Marie Griffiths et al., *InterConnections: The IMLS National Study on the Use of Libraries, Museums and the Internet Conclusions Summary* (February 2008), 4. The study found that while the majority of respondents continue to visit museums and libraries in person, the use of one resource type (like an online collection) often leads people to seek out additional information (such as in the form of an in-person visit).

A companion issue raised by those who resisted digitization early on was the idea of theft from archival collections. This line of thought held that once digitized materials were available online, it would be easy for someone to download the files, print them, and sell them, either as a book or as individual images – all without profiting the digitizing entity one cent. While it is possible that this may have occurred in isolated examples, it is also likely that such acts are the exception and not the norm. The likelihood of being able to pass off another institution's resources as one's own, and the ability to recognize a profit from doing so, are low enough as to make trying such a scheme very unattractive. Recent techniques such as using lower-quality images for online access, watermarking, or restricting access to collections are routinely implemented to help protect online assets.

Ironically, the fear of theft of library resources was a major concern during another period of great change in the realm of information delivery. In the 1890s, the idea of "open access" to shelves where books were stored – as opposed to submitting a request to a librarian, who then procured the book from "closed storage" – was slow to catch on owing to concerns of sticky-fingered patrons absconding with a library's leather-bound treasures. As Abigail Van Slyck recounts in her book *Free to All: Carnegie Libraries and American Culture 1890-1920*, a few pioneering institutions helped ease the librarians' collective fears.

"Although the question of free access to books was still hotly debated in the 1890s, large urban libraries joined the ranks of those institutions experimenting with open shelves. Of particular note was the Cleveland [Ohio] Public Library, which offered unrestricted access to all books at all hours beginning in 1890. Allaying fears about the wholesale theft of books, librarian William Howard Brett reported to the ALA that the practice had served to increase the library's circulation."

As more and more institutions saw the benefits of allowing patrons greater access to information, the "democratization of knowledge" began its unstoppable march across the country, aided by the establishment of Carnegie libraries in communities large and small. Similarly, with every

new collection that goes online and avoids falling prey to such piracy, this potential barrier to digitization continues to weaken, and researchers around the world are the ultimate beneficiaries.

Another considerable roadblock to digitization projects has historically been finding the staff to carry them out. When the technology was first making its way into the market, many commercially available scanners featured less-than-intuitive interfaces, and their relatively large size could impact their physical placement within a work setting. But even if a museum were able to purchase and house a scanner, it was usually finding someone to use it that presented the greatest challenge.

A common refrain heard from smaller institutions was often, “I’d like to digitize our collection of early photos of our town, but no one on our staff has the time to do it – we’re all too busy.” Operating most scanners is not difficult *per se*, but it does require significant staff time in order to be successful. Becoming a proficient and effective scanning professional is a situation that develops over time, and the time it takes to learn how to tackle complex scanning situations is often a luxury a museum of modest size cannot afford.

While operating complex scanners such as high-speed book scanners or machines that use proprietary software can require modest amounts of training time, most employees can become comfortable using desktop flatbed scanners in a relatively short time.³ Adding basic scanning duties to the job responsibilities of an existing employee should therefore be a matter of supervisory discretion and perceived competence on the employee’s part. In many cases, the greatest obstacle to completing a digitization project is not finding a designated amount of time in someone’s day to work on it; it is the simple fact that the process is not begun at all.

³ Penelope Kelly, *Managing digitization projects in a small museum*, Master’s project, University of Oregon, 2005. Accessed February 8, 2010. Available from <http://hdl.handle.net/1794/937>

Current challenges

While barriers such as cost, technology, institutional mindset, and staffing were early and frequently cited difficulties for museums looking to launch a digitization project, they have largely been addressed thanks to improved technology (which leads to lower costs), an increased comfort level with the process, and museum administrators' willingness to integrate scanning duties into existing job descriptions. However, a host of new challenges has arisen in recent years, as digitizing institutions make the transition from a mass digitization mindset to one of creating relevant, useful, and instructive experiences for online content consumers.

Scanning is easy; metadata is difficult

As digitization of large collections became easier with the rise in cheap, fast scanners, the volumes of digitized materials brought with them a new challenge: assigning accurate, useful metadata to digital records. "Metadata" is usually defined as "data about data." According to NISO, a non-profit organization that develops and publishes standards related to information management, there are three main types of metadata:

Descriptive metadata describes a resource for purposes such as discovery and identification. It can include elements such as title, abstract, author, and keywords.

Structural metadata indicates how compound objects are put together, for example, how pages are ordered to form chapters.

Administrative metadata provides information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it. There are several subsets of administrative data; two that sometimes are listed as separate metadata

types are *Rights management metadata*, which deals with intellectual property rights, and *Preservation metadata*, which contains information needed to archive and preserve a resource.⁴

Metadata can be configured in a number of different element sets, commonly referred to as schema. Common versions include Dublin Core (a basic set of 15 core elements and one of the most popular schema), Text Encoding Initiative (TEI), Metadata Encoding and Transmission Standard (METS) and more. Each schema offers differing levels of granularity, that is, the degree to which a record may contain detail about the object, from basic to highly sophisticated (See Figure 1).

⁴ *Understanding Metadata*, National Information Standards Organization (Bethesda, MD: NISO Press, 2004), 3.

home : advanced search : preferences : my favorites : about : help

Baylor University Library Digital Collections

menu off add document to favorites : add page to favorites : reference url

Search this object:

0 hit(s) : previous hit : next hit

View: document description

previous page : next page

Annual report of missionary of James Bartley

- 21moodybapt_04_59_001
- 21moodybapt_04_59_002
- checksums

Title Annual report of missionary of James Bartley

Creator James Bartley

Date 1956-1957

Geographical Location Uruguay

Material Type Annual report

Physical Description Carbon copy of letter text on 8.5" x 11" paper

Descriptive Note Two-hole punch marks

Keywords Annual report -- James Bartley -- 1956 -- 1957 -- Uruguay -- Missionary

Language English

Archival Identifier Folder 04, item 59

File Identifier 21moodybapt_04_59

powered by CONTENTdm © | contact us

home : browse : advanced search : preferences : my favorites : about : help : [Collection Home] login

Selections from the Texas Collection

menu off add document to favorites : add page to favorites : reference url back to results : previous : next

Search this object:

0 hit(s) : previous hit : next hit

View: document description

previous page : next page

Catalogue of Officers and Students of Baylor University 1868-1869 (Male Department)

- Page 1
- Page 2
- Page 3
- Page 4
- Page 5
- Page 6
- Page 7
- Page 8
- Page 9
- Page 10
- Page 11
- Page 12
- Page 13
- Page 14

Title Catalogue of Officers and Students of Baylor University 1868-1869 (Male Department)

Title Series Baylor at Independence Catalogues

Creator Name Baylor University

Creator Role Author

Creator Information Independence, Washington County, Texas

Publisher Name John F. Weishampel, Jr., Bookseller and Stationer

Publisher - Place of Baltimore, MD

Publisher Information No. 8, Under the Eutaw House

Date - Original Resource May 01 1868;

Creation

Language English

Description - Physical printed text on 4.5" x 7.75" paper -- cover shows evidence of water staining and is bound tightly into gutter -- cover bears stamp from Baylor Library reading, "Book; Pamphlet No. 2547" -- interior pages in good condition with some notations written in ink -- [24 pages]

Description - Content Catalogue of Baylor University for the 1868-1869 school year for the Male Department. Contains information about the Board of Trustees, university officers, students enrolled, curriculum, and "laws and rules of discipline."

Keywords Baylor University Independence 1868 1869

Coverage - Place Independence, Texas

Coverage - Date Range 1868-1869

Collection The Texas Collection -- University Catalogs

Institution Baylor University |

Rights Management Copyright The Texas Collection, Baylor University

CONTENTdm file name 459.cpd

powered by CONTENTdm © | contact us ^ to top

Figure 1: Granularity in metadata records can be described as less granular (top) and more granular (bottom). The more detailed a record, the more granular it is said to be.

Including accurate metadata with each online record is essential for the resource discovery element of a digital collection. If accurate, adequate metadata is not added to each item, it will be very difficult for search engines like Google to find the items and make them available to users. Metadata also ensures that items in your museum's collections can be virtually "tied together" through services that harvest metadata and associate disparate items with one another. One such example is OAIster, a catalog created by OCLC that harvests metadata from millions of records and makes them available through a central search via WorldCat.org. For Texas-based institutions, the Texas Heritage Digitization Initiative (THDI) offers resources for the creation and maximization of metadata for online collections, as well as information on best practices and strategies for improving end-user experiences.⁵

Another important purpose served by metadata is its ability to aid in digital asset preservation. The information associated with digitized items is fragile; as data is migrated forward from existing sources and formats into those used in future iterations, it is possible for data to be corrupted, misplaced, or altogether lost. Adding rich administrative metadata to an item will make it more likely that digitized resources created today will be useful to researchers 10, 50, or 100 years from now. Fortunately, most digitization software solutions automatically embed much of this information with each scan, and all database management solutions – like CONTENTdm or Greenstone – allow users to add even more fields.

As the preceding discussions illuminate, adding metadata to digital objects is not only important, it is essential to the long-term viability of all digital resources. While some metadata fields are harvested automatically (like titles, authors, and publication dates), other fields – notably those associated with

⁵The Texas Heritage Digitization Initiative Web site. Texas State Library and Archives Commission, Austin, TX. Accessed February 8, 2010. Available at <http://www.tsl.state.tx.us/thdi/>

full-text transcriptions, detailed descriptions, and page-level metadata – may require painstaking entry of information into a database, a process that can be time-consuming and is subject to any number of errors on the part of the staff member tasked with entering the data into the system. The importance of adding useful metadata to digital objects is such that it must be fully embraced by museum administration before any digitization project is undertaken. Online digital objects that have no useful metadata assigned to them may be likened to a turtle on a fence post: they are not likely to be found, no one knows how they got there, and no one quite knows what to do with them if they do find them.

Long-term storage and format migration

As vexing an issue as metadata is the two-pronged quandary of how to provide long-term storage and ensure format migration. Once a digital object is created, the creating institution must be prepared to preserve it for the foreseeable future. Taking the long view means purchasing the proper amount of storage, whether in the form of on-site resources or off-site repositories, with an eye to keeping resources safe and secure indefinitely.

The best types of storage utilize “spinning discs,” more commonly called hard drives. These can range from 1 TB external storage drives (like the one used in this project) to multiple terabytes of storage in the form of a local storage server. However, these solutions can cost thousands of dollars, and their installation and ongoing maintenance may require someone on staff with experience in programming and system administration. For this reason, many small museums may be tempted to use CD-ROMs or DVD-ROMs for backing up their digital collections. While a better solution than not backing up at all, these media are not considered archival, in that they require a device to be read, and they are prone to loss, scratching, or other kinds of damage. Faced with the choice between cost and potential

failure rates, however, many institutions may opt for the cheaper alternative, a route that could lead to data loss much sooner than most realize.

Related to the problem of long-term storage is the question of preservation formats. To newcomers in the digital collections realm, it may seem that there are almost as many file formats available as there are digital objects. Surveying the field of file types does little to change this impression: from GIFs to TIFFs, from JPGs to MPGs, the plethora of filenames and platforms can become quite daunting. While some formats have settled out to be the industry standards – for example, TIFF files scanned at 300 dpi are considered to be appropriate for preservation copies – there is still enough debate among the digitizing community that it can be a turn-off for institutions that are just getting started.⁶

Of course, settling on a file type is not enough. A digitizing institution must also be prepared to convert or “migrate” those files to future formats when and if the original format falls out of favor or is no longer supported by a majority of users. This is the truly difficult part of digital preservation, because despite experts’ best efforts to project what formats will be available in 10, 20 or 100 years, surprises occur with alarming frequency. For example, it was not long ago that PDF format was considered too proprietary and little-used to be a smart preservation format. Today, thanks to its growth throughout

⁶ To take just two examples, the Hudson River Valley Heritage Digital Service initiative requires master images be scanned at 48 bit color in 400 to 800 dpi, while the Illinois State Library Digital Imaging Program adheres to a 24-36 bit color format of 300-600 dpi for archival images. Hudson River Valley Heritage. Highland, NY. Accessed February 8, 2010. Available from <http://www.hrvh.org/about/digitalimaging.htm>. Cyber Drive Illinois. Illinois State Library. Springfield, IL. Accessed February 8, 2010. Available from [http://www.cyberdriveillinois.com/departments/library/who we are/bestpractices.html](http://www.cyberdriveillinois.com/departments/library/who_we_are/bestpractices.html).

the computing world, many institutions retain PDF/A (or “archival PDF”) versions of their digitized files as preservation or near-term access copies.⁷

While smaller museums can benefit from the trailblazing work being done by larger institutions in the areas of storage and file migration, they can also become paralyzed with indecision while waiting to see which solution comes out on top. This uncertainty to act is another example of the kinds of challenges facing museums struggling over the feasibility of starting a digitization project.

Providing access

Another major challenge of contemporary digitization projects is the question of providing access. In this case, the difficulty comes not from facing a lack of options; on the contrary, there are a number of alternatives available to fit the needs and budget of museums of every size. The challenges of access fall into two categories: providing an intuitive interface and clearing up questions of copyright.

An intuitive interface is crucial to a positive user experience, and it is important to make that impression quickly. Canadian researchers found that Web users can make judgments on the aesthetic quality of a Web site in 1/20 of a second. Thus, a positive first impression is key to drawing users in. But the same study showed that a “halo effect” also applies to Web users. In effect, if users believe a Web site is visually appealing, “then this positive quality will spread to other areas, such as the website’s content”.⁸

⁷ Sustainability of Digital Formats, Planning for Library of Congress Collections, *PDF/A-1, PDF for Long-Term Preservation, Use of PDF 1.4*. Washington, DC. Accessed February 10, 2010. Available from <http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml>.

⁸ “First impressions count for web,” BBC News Web site. January 16, 2006, Monday. Accessed January 30, 2010. Available from <http://news.bbc.co.uk/2/hi/technology/4616700.stm>.

But an online system must be more than just easy on the eyes. It must also feature elements such as easy-to-use navigation, quick load times, and, where possible, customizable content. When museums create digital collections that embrace these elements, users are more likely to spend longer amounts of time with the collection, thereby gaining a deeper appreciation of the content.⁹ But navigating the seemingly endless number of options – from homegrown systems to out-of-the-box solutions to Web-hosted services – can come to dominate museum staff members’ time and drain enthusiasm for the project. Throughout this report, I will explore several different options for presenting digital collections online, with a focus on meeting the aesthetic and functional needs of a small to medium-sized museum.

Putting the materials online is only half the problem, however. The questions of copyright and ownership also come into play, especially with materials that were produced after 1923. In general, legal scholars and information specialists agree that most items produced before 1923, with the exception of some unpublished manuscripts and so-called “orphaned works” (i.e. works where authorship and original copyright are difficult to establish), are safe for presenting in an online collection. These out-of-copyright works are considered to be in the public domain, so digitizing them and making them freely available poses little risk to the digitizing institution.¹⁰

Things become more complicated with materials published after 1923. Depending on format, authorship, and the terms in place at the time of creation, materials may fall into the public domain, they may be protected under the terms of their current copyright holder, or they may be caught in a “gray

⁹ Chris Norris, “It’s Not the Size of the Online Collection ...”, *Museum*, January/February 2010. Accessed February 12, 2010. Available from <http://aam-us.org/pubs/mn/onlinecollection.cfm>.

¹⁰ For a good overview of how to determine copyright status, see the Michigan Library Consortium’s “Copyright and Digitization” site. Lansing, MI. Accessed January 20, 2010. Available from http://mlc.lib.mi.us/cms/sitem.cfm/library_tools/copyright_/copyrightdigitize/

area” where it is difficult to ascertain who owns the copyright. While libraries and archives are afforded some unique protections under Section 108 of the U.S. Copyright Code, museums may find it more difficult to claim such exemptions for materials they wish to digitize. For most museums, the best advice may be: when in doubt, consult a lawyer.¹¹

Providing context and relevance

Another major challenge associated with digitized content is how to make the material relevant. A museum may post hundreds of digital objects on its site, but without relevance and context, they are little more than a laundry list of links. Assigning rich metadata is a good first step, but it takes a curatorial approach to digital collections to truly make them come alive.

Curators are essential to the health and utility of physical collections, and their expertise is no less important in the digital realm. A good curator will spend hours researching the backgrounds, settings, and cultural importance of the items in their collections, and the application of this information leads to exhibits that are rich in detail and context. But this process requires time and expertise, and in an institution where a single curator may be responsible for a very large collection with little or no assistance, adding the curation of digital collections to an already overwhelming workload may mean the online exhibit is largely ignored in favor of more pressing physical needs.

Curators may be forgiven for believing that merely placing items online will ensure their discovery and use by researchers. In reality, creating context could be argued to be even more important in the digital environment, simply because the opportunities for misuse are much greater. An

¹¹ William Maher, “Copyright: The Archivist and the Law” workshop, sponsored by the Society of American Archivists, held at Texas Christian University, November 5-6, 2009.

item viewed from a digital collection that is divorced from its context may be disseminated through unintended channels to users whose experience with the item is entirely uninformed.

For example, a digital collection of sheet music from the late 19th century, when casual racism and anti-Semitic sentiment were common, may contain an introductory narrative that places them in the proper chronological and sociological contexts. But if an image is removed from those contexts and sent via email or social networking sites to others with only the phrase, “Look at the racist stuff they’re showing at the XYZ Museum,” great damage can be done to a museum’s most important asset: its reputation.¹²

Museums and digitization

Museums can utilize a digitization program to meet two major goals: exhibition and preservation of content. These goals are not mutually exclusive, as the process of digitization can be tailored in such a way that both needs are met over the course of the project.

Online exhibits

The creation of an online exhibit helps extend the reach of the museum beyond its walls to a worldwide audience. The exhibits may be designed to stand alone permanently, they may compliment an existing exhibit, or they may be a temporary addition to the museum’s overall exhibition strategy. Because of the many available options for museums hoping to create an online exhibit, museums of all sizes and backgrounds have created them, from small local museums to powerhouses like The Metropolitan Museum of Art. They may be sophisticated or simple, multimedia-based or basic, but

¹² For an exploration of how so-called “digital object identifiers” may be utilized to help prevent this problem, see Lloyd A. Davidson et al., “Digital Object Identifiers and Their Role in Implementation of Electronic Publishing”, IEEE 1998. Accessed February 18, 2010. Available from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00730709>

regardless the execution, these online exhibits are important ways museums can advance their educational mission in an entirely virtual setting.

In many ways, museums could be said to have led the way toward how content is presented online. For years, museums have created interactive exhibits based around multimedia elements such as video, audio, photographs, and projected images. These elements have made their way into the online realm, as institutions from major corporations to local libraries utilize them to create engaging, informative online presentations. While they will never take the place of in-house interactives, museums have rightly seen that online exhibits can help reinforce information presented to traditional visitors while allowing access to a potential audience of millions who may never make an on-site visit.

One example of a museum that saw the value of adopting an online presence early in the digitization era is the Guggenheim Museum. Since 2001, their Web site (www.guggenheimcollection.org) has featured close to 200 items (described as a “selection of highlights”) from its permanent collection, all of which are searchable through a custom database. The Guggenheim notes that the site “also includes additional scholarly and contextual information, such as artist biographies, definitions of art-historical terms, concepts on art, and suggested readings,” all of which are designed to provide users from around the world with a sense of the Guggenheim experience, even if they never set foot in the museum’s doorway.¹³

Preservation

In this context, the idea of using digitization to preserve materials does not refer to the preservation of the physical asset. Here, preservation means preserving the *information* contained

¹³ Guggenheim Museum, Selections from the Collection. New York, NY. Accessed November 20, 2009. Available from (www.guggenheimcollection.org).

within the physical asset, including its coloration, size, printed information, physical appearance, and more. Using digitization as part of the preservation process adds a layer of protection for the physical asset, and it provides a measure of peace of mind should the unthinkable happen and the physical object be lost.

Museums that digitize their collections may find that they discover better solutions to the physical preservation challenges they face. For example, an archival collection may contain a number of early 19th century letters and manuscripts from a prominent local family. These items are often requested by researchers and genealogists, and years of use have rendered them brittle, stained, and torn. When the museum selects this collection for digitization, once the digital versions are created, the physical objects may be conserved and placed in permanent storage where they are no longer accessible to the outside public. Reprints or online versions can provide the content-based information being sought by most researchers, with the added benefit of portability: if a researcher wants a copy of a letter, it is a simple matter to create one.

Not explored fully in this project but worth mentioning here is the possibility of monetizing archival holdings through digitization. Many museums make a tidy sum by selling authorized reprints of items from their collections in their gift shops. Avoiding for now the potential ethical questions raised by such a practice, it is worth noting that once a museum has created a preservation-quality digital image of a physical object, it is only necessary to reformat it and print the requisite number of copies in order to turn that historic resource into a contemporary moneymaker.¹⁴

¹⁴ As Debra LaKind, Head of Rights, Licensing and Visual Archives at the Museum of Fine Arts, Boston told Epson: “The additional revenue stream generated by the sale of fine art prints allows us an opportunity to provide increased digital access to the collection.” “Focal Points – Fine Art Reproduction.” Epson America. Long Beach, CA. Accessed January 28, 2010. Available from http://www.epson.com/cgi-bin/Store/SuccessStories/ssDetail.jsp?title=ss_mfaBoston&detailType=fineart

Best practices for digitization projects

Although the term “best practices” may be a bit of a misnomer – one author points out that because of the rapid evolution in the world of digitization, the word “guidelines” may be a better choice (Liu, 1) – it is important to identify some basic standards that have emerged which provide a good starting point for museums looking to start their own digitization project.

1. **Determine relevance:** Many digitization projects are begun without anyone asking the very important question, “What do we want this collection to *do*?” Before a museum sets off on an ambitious plan to scan their entire archival collection, a team consisting of the curatorial staff, the education department, the administration, and the marketing/public relations staff should be created to assess and assign priorities to any potential digitization projects before any items are placed on a scanner.
2. **Identify key staff and get the right training:** Every digitization project is different, so it is important that every museum determine who will be doing the digitization, and what kinds of training they need in order to successfully carry it out. Is the museum going to use volunteer labor, or will the project be part of a staff member’s job description? Can we train digitizers in-house, or is a trip to a 2-day seminar in order? The answers to these questions may help determine the project’s direction and its workflow.
3. **Decide on a metadata schema:** This is crucial, as it is one of the major decisions that determines if a project will succeed or fail. Defining the relevant types and amounts of information required

for each digital object must be established early on in the process, or else the museum risks creating incomplete and ineffective digital records. The type of schema chosen is not as important as the commitment to consistent, accurate adherence to it through all stages of the digitization process.

4. Set technical standards and stick to them: It is common practice to scan items in 24-bit color at 300 dpi for preservation files; access copies that become part of the online collection may be adjusted to fit the museum's particular needs. TIFF format is best, with JPG2000 another good option for images; for video, MPG4 is a common standard, and uncompressed FLAC files are often used for audio files. It is also a good idea to run checksums – a process that creates a digital “fingerprint” for every image for later tracking and integrity checking – on all files in order to determine if any data loss has occurred during transfers from machine to machine.¹⁵
5. Plan for triple redundancy in backups: A single copy of a digital file is prone to loss, corruption, or unintended alteration, so it is a good idea to plan to keep at least 3 copies of every file. Typically, that would equate to one file on the local machine (the computer attached to the scanner, for example), an on-site backup (perhaps on an external hard drive or a local server) and a third kept off-site. (Current best practices suggest these off-site backups are most effective when housed at least 100 miles away from the digitizing institution. The thinking is that in the event of a natural disaster, a site 100 miles away is a sufficient buffer to provide safety of the digital assets.)

¹⁵ For a good starting point to find best practices for digitization projects, see Yan Quan Liu's "Best practices, standards and techniques for digitizing library materials: a snapshot of library digitization practices in the USA", *Online Information Review*, Vo. 28, Issue 5, pp. 338-345, 2004. Available from <http://www.emeraldinsight.com/Insight/viewContentItem.do;jsessionid=E6AC16AA8A2F1D034A952C4864323333?contentType=Article&contentId=862266>

Adhering to these five guidelines will help provide a solid background for a digitization project of any size and scope. While subtle gradations are common from project to project, these fundamentals are almost uniformly consistent, and they will help museums of any size find success in their digitization projects.

CHAPTER TWO:

Selecting a Museum Partner

As one of the main thrusts of this project became clear – the decision to digitize and place online items from the archival collection of a small to mid-sized museum – it became necessary to select a museum with which to partner. I was also hoping to do some survey research regarding how these museums viewed the process of digitization, and whether any had undertaken a digitization project of their own. I combined these two needs into one instrument: an online survey I created using Survey Monkey (www.surveymonkey.com), a Web-based survey creation site that offers both a free and fee-based surveys for businesses, non-profit organizations, and individuals.

After signing up for a “pro” account, which provided a greater level of sophistication and data collection than the free account, I created a survey that consisted of 20 questions. The survey consisted of simple “yes or no” questions, multiple choice questions, and questions that used a Likert scale to determine the level at which respondents rated their agreement with certain statements (from “not at all” to “agree completely,” for example).

Methodology

Potential respondents were chosen from a number of sources, but primarily from a listing of museums that had provided their contact information to the Texas Association of Museums (TAM) for inclusion in the MuseSearch section of the “Find a Museum” tab on their Web site (<http://www.io.com/~tam/index.html>). After calling to verify that the email addresses listed on the TAM site were correct, I sent a link to the directors (or, in the absence of a director, the curator) of 12

potential museums from across the state. Also, TAM executive director Ruth Ann Rugg graciously agreed to place a link and a brief explanatory paragraph in two successive issues of TAM's electronic newsletter. The newsletter is sent to all TAM member institutions, with an approximate circulation of 2,000 individuals per month. In all, 10 respondents (representing 10 unique institutions) completed the survey, although 14 individuals from unique institutions answered at least one question. For a full list of questions and complete survey responses, please see the Appendix.

Survey results

Digitization and general knowledge

The first five questions dealt with respondents' levels of general knowledge about digitization. Of the fourteen people who answered the first question, it was determined that all fourteen were at least familiar with the term "digitization" as it applied to museum archival collections. A majority (64.4%) rated their knowledge of "the theory and outcomes of the digitization process" as "some," while 28.6% rated their familiarity as "very familiar." Just over 7% answered that they had no knowledge on this question. A combined 14.2% answered that they had little or no knowledge of the steps involved in a digitization project, but a majority (85.7%) answered that they were either "somewhat" or "very familiar" with them. Of the fourteen respondents completing question five, only one indicated that their institution had never undertaken a digitization project. The majority of those responding in the affirmative had digitized photographs (ten of fourteen, or 71.4%). Letters, three-dimensional objects and books rated the next highest response rates.

Institutional holdings

Questions in the next section sought to determine the size, scope and condition of the items in respondents' collections. Nine of twelve respondents (75%) indicated that their archival holdings

contained more than 1,000 pieces. The next question was multiple choice in format and allowed respondents to list the types of items held in their archival collections; twelve of fourteen total survey takers responded to this question. All twelve indicated loose (or unbound) photographs in their collections, with “pamphlets, brochures and promotional materials” ranking ten responses, followed by letters (nine), official documents (nine) and other books (nine). When asked to rate the condition of the items in their archives, a majority (six of twelve) answered “good,” but a combined 33.3% answered either “fair” or “poor.”

Willingness to undertake a digitization project

The next section comprises the crux of the survey instrument insofar as the answers to the next four questions (nine through twelve) serving to identify which institution (if any) would be willing to partner with me for this project. Twelve of fourteen total survey takers answered question nine, which sought to gauge respondents’ reaction to such items as cost of digitizing items, ownership of preservation-quality images, and hosting materials in an online environment. All twelve respondents indicated their willingness to participate in a digitization project if the process were free, but three of the twelve also said they would do so if it required a fee. Three respondents also indicated they would be willing to send their items no more than 100 miles away in order to be digitized by trained staff on specialized equipment.

The subject of ownership was addressed in the next section. Only two of the twelve respondents said they would allow the digitizing entity (in this case, Baylor University) to add digitized materials from the museum’s collection in its own online collections. However, five of the twelve were willing to allow Baylor to keep copies of the digitized items for preservation, and eight of the twelve were willing to

allow Baylor to keep a copy for preservation if Baylor offered to provide access-level images back to the museum for hosting on their own Web sites.

The next two questions (ten and eleven) do much to reveal the respondents' overall feelings toward two key aspects of this project: collaboration and online exhibits. While a combined 41.7% of respondents were either not at all interested or only somewhat interested in taking part in a "collaborative, Central Texas-focused digitization project," a majority (58.3%) were either "interested" or "very interested." But when asked in question eleven if respondents were interested in learning more about how to host digitized materials on their own Web sites, a very high 91.7% (or eleven of twelve respondents) answered in the affirmative.

The final major question of the survey, question twelve, would help me narrow my focus to one particular institution whose attitudes, collections, and willingness to collaborate would point to them as a natural fit for this project. Two-thirds of respondents to this question (eight of twelve) indicated that they would be interested in being part of a pilot project that would highlight a collaboration between myself (representing a "major institution with specialized training and equipment") and an area museum. Of those eight surveys, seven provided follow-up contact and demographic information on which I could base my decision. All that remained was to determine a list of several potential candidates for partnership, initiate follow-up contact, and establish a working relationship.

Partnering with the Railroad and Heritage Museum

Of the potential museums that had indicated a willingness to partner with me for this project, the most likely prospect was the Railroad and Heritage Museum (RHM). Located in a restored Santa Fe depot in downtown Temple, Texas, the RHM features a large collection of railroad-related items ranging from timetables and brakemen's lanterns to rolling stock (i.e. locomotives, Pullman cars, freight cars,

etc.) and an extensive collection of archival holdings. I contacted Craig Ordner, the museum's archivist, to inform him that I had chosen the RHM from a list of contenders, and he reacted very positively. I arranged a meeting at the RHM for September 11, 2009 with Ordner and executive director Judy Covington to discuss the project.

At the initial meeting with Ordner and Covington, I outlined how I saw the project working. Drawing on Ordner's experience as archivist, I asked the RHM to identify a collection of approximately 100 items that they had hoped to digitized but had so far been unable to accomplish, whether for lack of time, resources, expertise or any other reason. These items would then be delivered to the Ray I. Riley Digitization Center at Baylor University, where I would scan them and store them on an external hard drive. Then, I would work with the museum to place the items online and accessible via their homepage. Both Ordner and Covington expressed great interest in the project, with Covington saying that she had hoped for some time to start a large digitization project, but the museum had so far been unable to do so.

As the meeting progressed, Ordner identified a likely collection for digitization: a collection consisting of railroad engineering maps, originally housed at the Houston Public Library. This collection had been donated to the museum by a collector in the Houston area who rescued them from the trash pile when the Southern Pacific Railroad's land office was planning to dispose of them. Drafted between the late 1800s and early 1900s, the maps show the rights-of-way of several major Texas railroads, including the Texas & New Orleans, the Galveston Harrisburg & Santa Fe, and the Missouri Kansas and Texas ("Katy"). The maps' size of 12" x 32" made them very difficult to scan on the RHM's flatbed scanner, and their frequent use by researchers made them an ideal candidate for digitization and online presentation.

The meeting concluded with an agreement between the museum and myself wherein RHM staff would deliver maps from the collection to Waco for digitization over the course of several months. This agreement represented the completion of the survey and initial setup phase of this project.

CHAPTER THREE:

The Digitization Process

The digitization process for the materials from the RHM's Houston Public Library collection of railroad survey maps took place at the Ray I. Riley Digitization Center (the RDC) at Baylor University. Located in the Moody Memorial Library, the RDC contains a number of highly specialized scanners, high-speed computers, and localized file storage solutions. The RDC digitizes materials for the Baylor library system, including its special libraries, as well as projects for the university community as a whole. As a member of the Digitization Projects Group of the Baylor Electronic Library, I was able to secure the use of RDC equipment to digitize the maps from this collection. Digitization of the maps took place after normal business hours and on weekends, so as to separate this project from ongoing RDC work, and to avoid any scheduling conflicts with RDC clientele.

Equipment and software

The Cruse CS-285 ST large format scanner

The scanning equipment used for this project consisted of a Cruse SLT-285 large format scanner, usually referred to by its maker name ("Cruse") or as the "large format scanner" (See Figure 2). The Cruse features a 59" x 88" vacuum table and a 36-bit color scanning camera that allows users to create items up to 10000 x 15000 pixels in size. Designed to scan large format items such as maps, large prints, and works of art on canvas, the Cruse at the RDC is one of only five in the state of Texas, and the only one available at a Texas university. The ability to use the Cruse was crucial to the success of this project, as the maps measure approximately 12" x 32" each, making them too large for even an oversized flatbed scanner. With the Cruse's ample space, I was able to scan 6-8 maps at one time, depending on

how they were arranged on the platform, and each map was imaged at a robust 300 dpi, the recommended resolution for preservation quality images.

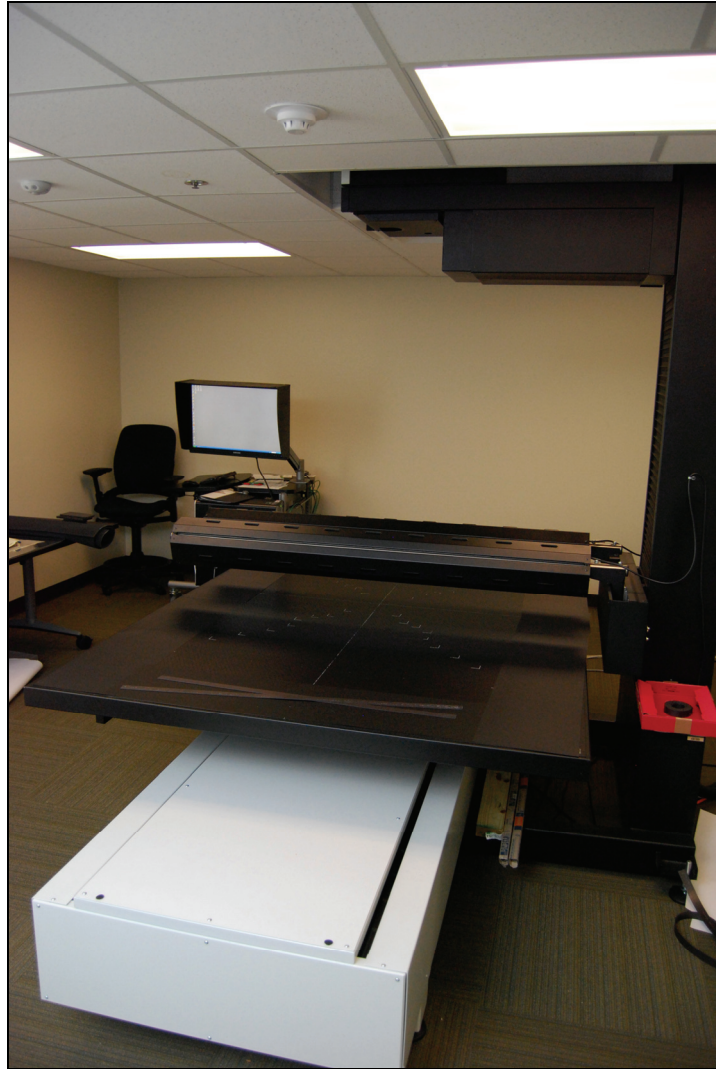


Figure 2: Cruse SLT-285 large format scanner

Western Digital “MyBook” external hard drive

Because the output files for each map (the preservation, highest-quality image and its online access derivative) would take up a great deal of space, the RHM purchased an external hard drive for

use on this project. The museum provided a 1 Terabyte (TB) “MyBook” external hard drive from Western Digital, which it purchased at a retail consumer electronics store for approximately \$100. At the close of scanning and digital derivative image creation, the total space consumed was 167 Gigabytes (GB), leaving a total of 763 GB free for future use. The hard drive was returned to the RHM and will serve as the primary storage device for the project, although the online derivatives are housed on Flickr’s servers, and a second set of backup copies are housed on the RDC’s preservation server.

Adobe Photoshop CS4

All image manipulation was undertaken in Adobe Photoshop CS4, the latest version of the software. Initial edits to the raw scans – including a first-pass crop and a light sharpening effect – were applied to the preservation quality images, and further re-sizing and color correction were applied to all access-level derivatives. These changes are considered “basic” level adjustments and could be completed by novice level Photoshop users.

Materials processing

The materials in this collection were delivered to the RDC on September 24, 2009 by executive director Judy Covington and archivist Craig Ordner. Unbeknownst to me, RHM staff had decided to bring the complete collection of 770 maps in the hopes that more could be scanned if time permitted, with the understanding that no more than 100 were expected. (Although previous discussions with the RHM had indicated that a sample of 100 maps would be appropriate for the project, I was able to scan close to 400 maps over the course of approximately three months; of those, 326 were included in the Flickr collection.) Ordner brought a standard outgoing loan agreement form for me to sign in order to receive the maps; I did so, acting as a borrower on my own without formal ties to Baylor University so as to insulate the university from potential liability in the event of loss or damage.

The maps were housed in large over-folders that broke the collection into more than fifty sub-folders of between five and thirty maps each. The over-folders were labeled by county name or railroad name, often down to the level of a particular branch or trunk line, common identifiers used by large railroads. Each folder was also assigned an identification number. I used these folder identification numbers and the accession numbers assigned by the RHM to each item as part of a custom identifier I created for the project. For example, a map from folder 53, with access number 1235 in the RHM archives, received an identifier of rhm_hpl_53_1235, which corresponds to Railroad and Heritage Museum (rhm), Houston Public Library collection (hpl), folder 53, item 1235. This assigned identifier helped me track the items through the workflow process, and provided an additional “tag” for the metadata associated with the maps that became part of the Flickr collection.

After the maps were assigned an identifier and logged in a master spreadsheet, the folders were housed in a locked, climate-controlled area of the RDC for the duration of the project. The maps were brought from this location to the room that houses the Cruse for scanning, then returned to the secure room at the close of each scanning session.

Scanning of the materials took place after normal business hours and on weekends for close to three months. As previously mentioned, the Cruse could accommodate between 5 and 7 maps at one time, depending on their size. For an average load of 6 maps, the scan time was approximately 20 minutes.

Following the scanning of a batch, I imported the files into Photoshop CS4, where I applied a basic crop to eliminate extraneous background pixels; applied a layer of the “unsharp mask” effect to eliminate the very minimal softness that is an artifact of the scanner; and saved the files to the external

hard drive. These slightly-modified files serve as the preservation or “master” files, in that they received the lowest levels of post-scanning processing.

Once the preservation masters were created, I opened them in Photoshop CS4 and created the access derivatives. This was accomplished by applying a closer crop on each image, ideally using an inner box that was drawn around each map as a guide. I then adjusted the color levels and saturation levels to create a more accurate depiction of what the maps would have looked like when they were originally drafted. These “access derivative” versions were then saved in a separate folder on the external hard drive for later uploading into Flickr (See Figure 3).

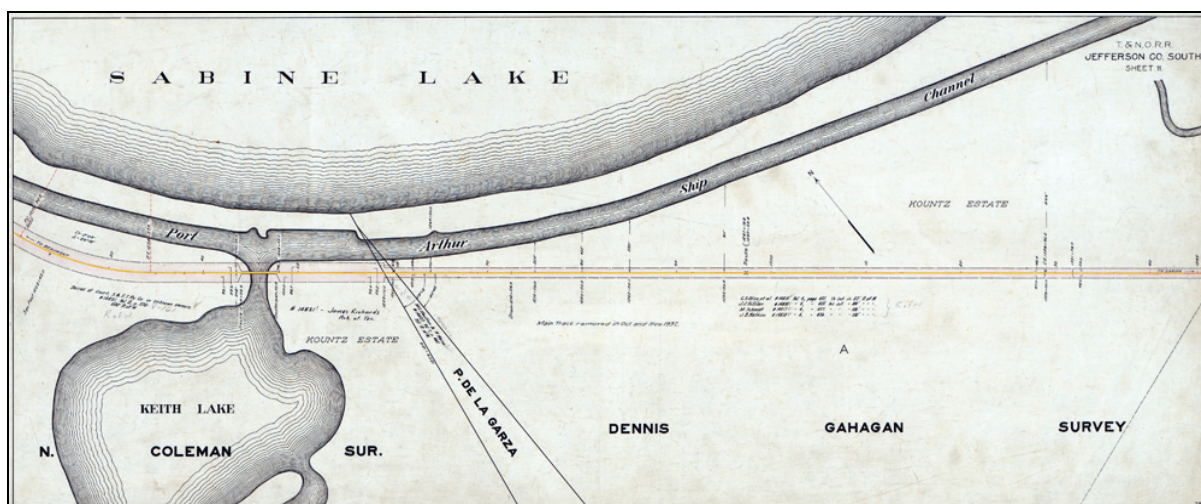
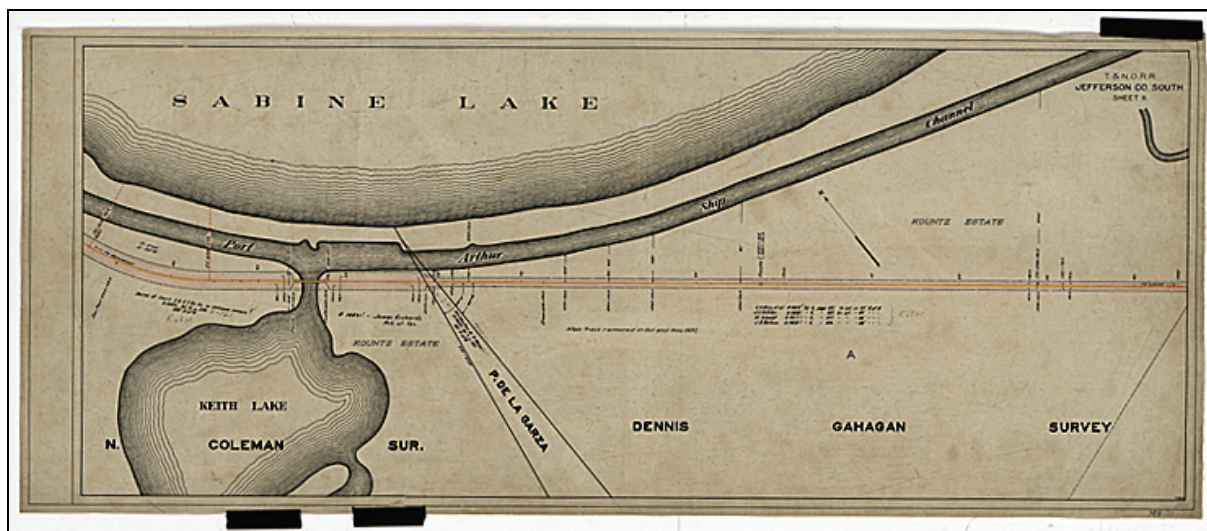


Figure 3: A sample preservation master image (top) and its access derivative (bottom). The black rectangles in the preservation image are magnets that were used to hold the maps flat on the Cruse scanner's platform during digitization. The versions available online were adjusted to provide better contrast for computer monitors.

As a further security measure, all files generated for this project were deposited on a local files server located in the Riley Digitization Center, called Dante. This secondary backup was further augmented by an off-site backup to the servers of the Texas Digital Library (TDL), of which Baylor University is a member. These two additional backups give double redundancy for the digital assets, providing at least two distinct opportunities for retrieval and restoration in the event of a catastrophic data loss.

Following this process, over the course of several months I was able to scan close to 400 maps, and after confirming that the digital files were backed up and all derivative files were successfully created, I returned the maps to the RHM on December 5, 2009.

CHAPTER FOUR:

Creating an Online Presence

The completion of the digitization and materials processing phase of the project led into the next step: placing the materials online. This process proved to be more challenging than I had originally anticipated, as it required the exploration (and subsequent dismissal) of several options before settling on the services that offered the right combination of features, ease of installation, ease of upkeep, and usability for the display, discovery, and exploration of the digital assets.

For this collection, I wanted to create two different online opportunities, a digital asset management system (DAMS), and an online exhibit. The differences between the two are as follows:

Digital asset management system (DAMS): Systems driven by a database in which material is organized, described, and arranged for retrieval. DAMS may be seen as a listing of resources sorted in some usable way, with each record containing information that may include an assigned identifier, a date created field, a description field, and other metadata. For researchers wishing to view all of the items digitized from this collection, a DAMS is the proper solution.

Online exhibit: A separate entity entirely from the DAMS, online exhibits are often more interactive, provide more contextual information, are inherently user-friendly, and serve as a gateway into a collection. For this project, the online exhibit would allow visitors to the RHM's Web site to learn more about the collection of maps, how they were created, the artistic nature behind them, and other information not found in a DAMS.

Because of the varied nature of the types of people who might be interested in accessing the digitized maps, I believe this two-pronged approach allows for the largest number of people to access

the maps in the greatest number of ways. For example, a dedicated Texas railroad researcher may be more interested in scanning through the hundreds of records available for the collection in search of a specific surveyor's name; for this user, the DAMS is the best option. But for a casual visitor to the RHM Web site, the online exhibit may prove to be an intriguing exploration of a topic on which they had no prior knowledge. From there, it may lead the user to peruse the DAMS, or the user may see the online exhibit as an enticement to visit the brick-and-mortar museum. These are just two examples of the myriad possible entry points into the digital collection, so having a DAMS and an online exhibit will likely provide the RHM with the greatest coverage for the collection.

The digital asset management system (DAMS)

Selecting a DAMS proved to be one of the more challenging aspects of this project. DAMS may be very complex or ineffectually simple; they may be robust or streamlined; they may be thousands of dollars or free, driven by open-source software; and they may require as much or as little technical expertise as one can muster. For these reasons alone, it was imperative that a DAMS be selected that would address the needs of the RHM's digital collections while not exceeding the capabilities of its staff.

It should be noted that the webmaster for the RHM Web site is a former member of the board of trustees, a retired pediatrician who is a self-professed "dabbler" with computers. He maintains the site as an unpaid service to the museum, and as a result, it is a very basic, table-based site. He has limited database systems management experience, and little or no software coding experience. Thus, it was imperative that any DAMS chosen would be easy to install, simple to navigate, and require little technical expertise to resolve any software-related issues. The RHM webmaster's novice level informed my search as I evaluated several potential DAMS.

Option one: Greenstone Digital Library Software (www.greenstone.org)

Greenstone is an open-source software for “building and distributing digital library collections” (www.greenstone.org). Created by the New Zealand Digital Library Project and distributed with help from UNESCO, Greenstone was created to provide a software solution for libraries, archives and museums around the world to create their own online digital libraries. Available in a number of languages, Greenstone has been implemented by institutions in countries around the world, including Afghanistan, France, Italy, and the United States.

Greenstone’s major advantage is its open-source nature. Open-source software is generally available free of charge and takes advantage of a worldwide community of users who create upgrades, patches and bug “fixes” so the software can be adapted to its greatest use. Because it is open-source, Greenstone allows institutions with little or no budget for online collections to acquire software that will help them build an online presence. Also, its Web site provides a wealth of information in the form of multi-lingual documentation, a wiki, example collections and helpful guides.

Greenstone’s biggest drawback – and the reason it was not chosen for this project – is the advanced level of computer coding and access to network resources required to install the software. Greenstone is not designed to be a “turn-key” solution; that is, where a simple install command and minimal configuration are all that are required to get the software up and running. To install Greenstone, one must have at least a basic understanding of HTML coding, as well as server-level access to the equipment running an institution’s Web site.

Because I would have to go through a volunteer (the former board member) in order to get this kind of access, it was determined that installing Greenstone would be difficult to impossible for this project. Also, the complex nature of the process needed to configure the collections for online display

would require an advanced knowledge of the software in order to make it feasible, and knowing that I would be unable to train anyone involved with the RHM software to become proficient in this process, I chose to abandon the pursuit of Greenstone for the DAMS.

Option two: CONTENTdm (www.contentdm.com)

CONTENTdm is a digital content management software that “handles the storage, management and delivery of ... digital collections to the Web” (www.oclc.org/contentdm/overview/default.htm). Created and administered by OCLC (a nonprofit computer library and research organization), CONTENTdm is a powerful database management system that can support thousands of digital objects, along with their metadata and relevant finding aids. The Electronic Library at Baylor University uses CONTENTdm to deliver its digital collections, and as such, it is the home to almost 50,000 unique items spanning dozens of collections.

CONTENTdm is open-source, but it is not free. This means that while there are numerous CONTENTdm users around the world who can provide input on how to solve particular programming problems, and while the software is adaptable to a certain extent, there are large parts of the code that are proprietary, meaning they cannot be altered by users. Also, there is a substantial cost to acquire CONTENTdm: \$30,000 per year for up to 50,000 objects, and \$50,000 per year for unlimited. Thus, while it provides fairly robust searchability, metadata support and display options, it is very expensive, needlessly complex and too much software for a project of this scope. For these reasons, plus the required amount of technical savvy needed to run it properly, I chose not to use CONTENTdm for this project.

Option three: Flickr (www.flickr.com)

The third option for the DAMS for this project was Flickr. Part of the Yahoo! network of Web sites, Flickr is a “photo management and sharing application” that aims to let users “make their content available to the people who matter to them” and to “enable new ways of organizing photos and video” (www.flickr.com/about/). A basic Flickr membership is free, and users have uploaded more than 4 billion photos and videos since Flickr’s launch in 2004.

The major advantages for using Flickr for this project are that it requires no installation, no coding experience and no programming expertise whatsoever. Users of all skill level can upload, describe and “tag” photos in Flickr. Because there is nothing to install, there are no adjustments necessary to the RHM’s Web servers. Likewise, because everything is accomplished using straight-forward, visual interfaces – from uploads to metadata to “tagging” – there is nothing to configure and no code to write. Flickr is open to the world, so once photos are uploaded, their associated metadata is harvested by Flickr’s search algorithms, and it is similarly added to search engines like Google and Yahoo! Another big advantage to using Flickr is that all the work needed to upload and describe the digital images of the RHM maps would be done through Flickr; once complete, a single URL could be provided to the webmaster of the RHM site for posting on the front page, along with text that would guide visitors to the Flickr collection. (For example, “Click here to view our online collection of early 20th century railroad maps!”)

Flickr does have a few disadvantages. While a basic account is free, in order to support the larger quality images associated with this project, it was necessary to upgrade to a “Pro” account for \$24.95 per year. The Pro account supports unlimited uploads of up to 10,000-pixel width across, a very

large format for Web viewing. (By contrast, the images in the RHM collection measure 5,000 pixels across the long side.)

Pro accounts also come with the option to track statistics for the account, including total number of unique views of each object, a quantitative measure that may prove useful to RHM staff as they describe how the Flickr collection is useful to the museum's visitors. Another minor disadvantage is that Flickr is not particularly flexible in terms of how items are displayed. One user's Flickr profile looks much like the others, with the exception of the items in the "photostream."

Overall, the advantages to using Flickr far outweighed the disadvantages, and after evaluating and rejecting Greenstone and CONTENTdm for this project, the decision was made to use Flickr for the DAMS for this project.

The utilization of Flickr

Because of the size of the access images created for the online collection (5000 pixels wide), it was necessary to sign up for a Pro Flickr account. I started by creating a generic email address (rharchives@yahoo.com) that would serve as the main contact email for the Pro account. Using the "gift" feature, I then used my personal Yahoo! account to sign up for a Pro account and "gifted" it to the RHM account, thus avoiding the need to associate my own credit card information with the RHM Flickr account (See Figure 4).

Once the Pro account was established, I used Flickr's batch uploading tool to upload the 326 items in chunks of 15-25 photos at a time. By default, the photos became part of my "photostream," a Flickr term for the collection of every photo or video uploaded by a user. Once the maps were part of my photostream, I created a collection entitled the "Houston Public Library railroad engineering survey

maps” collection. This collection was further sub-divided into 31 sets based on how the maps were grouped in their over-folders. As the RHM moves forward with its Flickr collection, it will be able to add more collections to its photostream. For example, it could create a collection of photos from events held at the museum, or of images scanned from other archival holdings.

Once the maps were organized, I began entering metadata for each item. Every item in Flickr has a number of customizable fields that can be edited to make the photos more useful. The two main fields I used are “Title” and “Description.” For each map, I assigned a useful title based on information found on each over-folder. The over-folders generally listed a Texas county of origin for the maps; a branch or division of the railroad; and the railroad’s name. In addition, each map was numbered by its creators, typically as “Sheet no. 1” or “Sheet 1.” So, each map’s Flickr title became a variant of County Name, Division/Branch, Railroad – Sheet Number. This will allow users to quickly search through title results using meaningful identifiers.

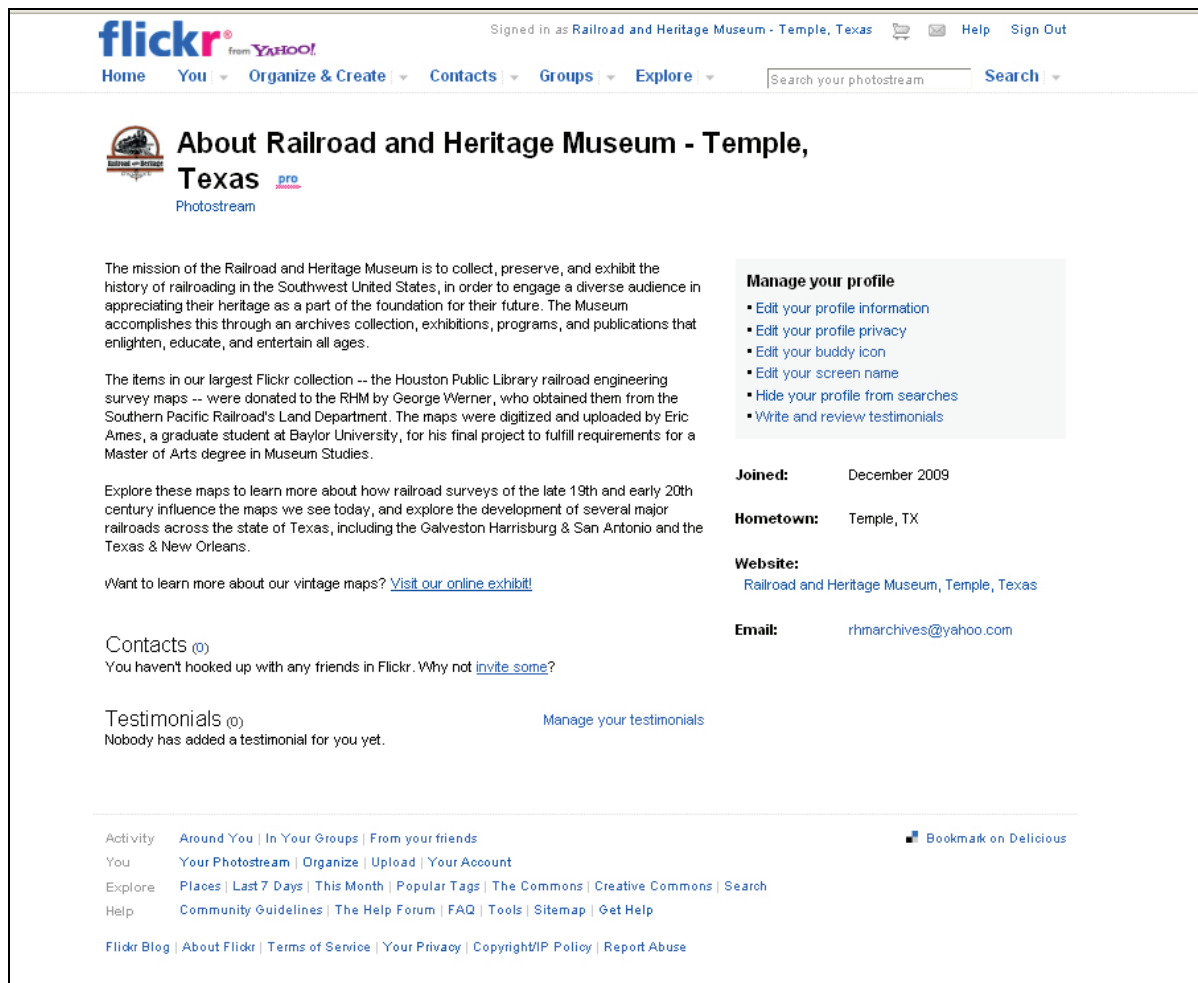


Figure 4: The RHM homepage in Flickr

For the Description section of each item, I created the following template:

Title -- Map of portions of <county name> County, Texas for the <railroad name> Railroad -- <Name(s)>, surveyors -- ink on treated linen paper -- <Project ID#> -- To order archival quality reprints, contact rhmarchives@yahoo.com

Using the metadata I had already harvested from the maps, I filled in the template for each item and placed this information in the Description box for each record. This standard format allows users to

quickly ascertain the major information associated with each map, and it standardizes the kinds of searchable information available for the records (See Figure 5).

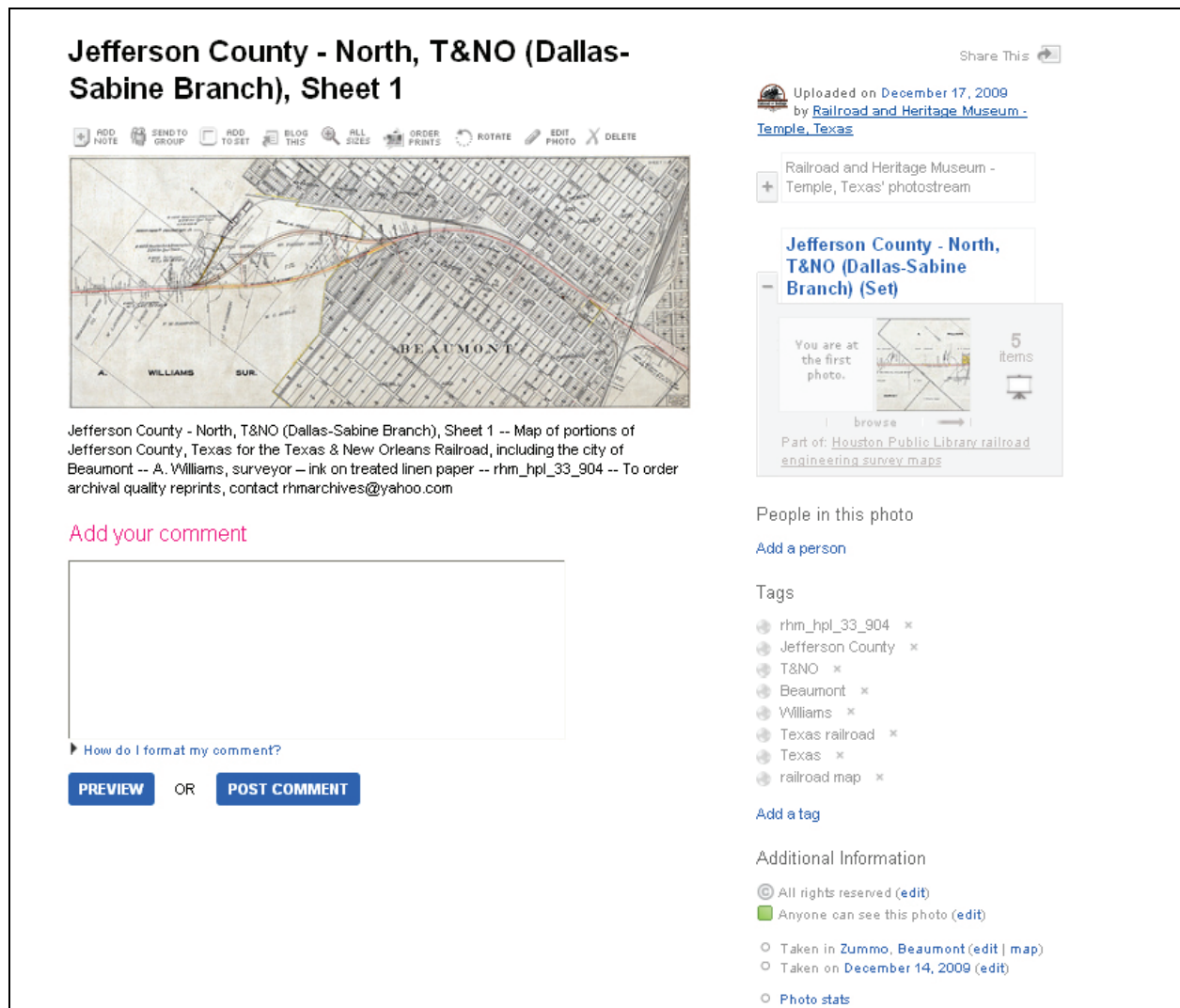


Figure 5: An item record view of a map from the collection. The caption below the image contains the standard description, while the “tags” (keywords) associated with the map are listed at bottom right.

The last step in creating the records was to “tag” each record. In Flickr’s terms, to “tag” means to assign keywords to each item. Tags then become a searchable option for users looking for certain

words in all of the photos available on Flickr. For example, a search for the tag “Waco” would bring up every photo uploaded by Flickr users who tagged the word “Waco” to their photo.

Each map in the Flickr collection is tagged with as much relevant information as possible. As a rule, I included its project identifier and county name, along with names of railroad lines, towns or cities, surveyors and other major elements where possible. For example, the photo of Jefferson County – North, T&NO (Dallas-Sabine Branch), Sheet 1 contains the following tags: rhm_hpl_33_904 (its identifier), Jefferson County, T&NO (the railroad), Beaumont (the town shown on the map), Williams (the names of a surveyor) and three generic tags: Texas railroad, Texas, and railroad map.

Including a relevant title, description, and keyword “tags” to each map will make them more findable by searchers both within Flickr and through Web search engines like Google and Yahoo! If they choose to do so, the RHM could also add tags like “railroad,” “heritage,” “museum” or “Temple, Texas” to each item in order to further increase the chances for the maps to be found by searchers.

An added feature offered to Flickr users is the option to “pin” photos to a point on a world map that indicates where the photo was taken (See Figure 6). In instances where a map contained the name of a town or city, I used the Flickr map function to correlate the area encompassed by the map with its appearance on the modern map. For example, the town of Sacul, in Nacogdoches County, has not changed much since its original platting in the early 1900s. Therefore, the railroad survey map and the modern Flickr map line up very closely to one another, allowing users to see how the town was laid out by the railroad in 1903, and how little that system of roads has changed in the intervening years.

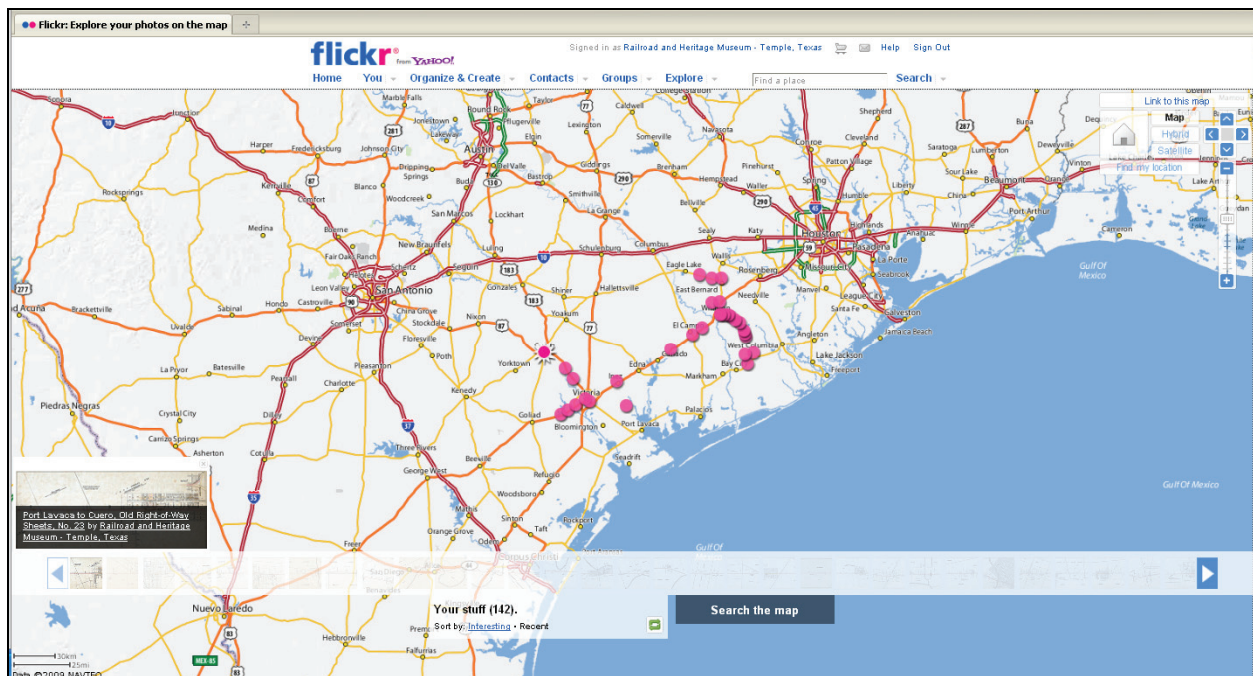


Figure 6: The Flickr map. The pink dots represent sites where maps are “pinned” to their modern day equivalent. The small window at lower left shows a thumbnail of a map that can be access by clicking on its associated pinpoint.

Overall, the process of uploading, describing, and making useful the digital items in Flickr was a smooth, intuitive, positive experience. The decision to use this tool, which is hosted off-site, supported by trained professionals, and reasonably priced, should provide a heightened measure of stability and usefulness for RHM staff moving forward with this and other digital projects.

The online exhibit

As previously noted, the online exhibit portion of this project differs greatly from its digital asset management system (DAMS) counterpart. The goal of this online exhibit is to provide context, meaning, relevance, and access to a diverse array of audiences through storytelling, a deeper examination of featured objects, and an easy-to-navigate structure.

While the DAMS allows for in-depth discovery of the digitized resources, the online exhibit was conceived as a way to engage visitors with the story behind the maps: why they were created, how they were used, how they arrived at the Railroad and Heritage Museum, and how they found their way onto the Internet. The execution for this part of the project would have to integrate images (a multimedia element), text, a zoom feature for enlarging objects, and a navigation system that was both simple and intuitive. At the outset, I examined three different approaches for the online exhibit: an HTML coded version, the Virtual Exhibit add-on for PastPerfect, and Pachyderm, a project of the New Media Consortium (NMC).

Option one: HTML coded version

Creating an online exhibit using HTML code was easily the most time-consuming and difficult option. On the plus side, HTML code is flexible, allowing users to create Web sites based on their specific needs. It can support multimedia elements, including audio and video, and pages created in HTML can be added to existing Web sites without the user having to purchase proprietary software solutions or relying on an outside vendor to host the pages.

However, creating Web pages in HTML requires extensive knowledge of the coding language for best results. Pages built in HTML require a good deal of foresight and planning if they are to be successful, and with the increasing number of template-driven, “plug and play” style Web site generators on the market today, people with expertise in HTML coding are increasingly more difficult to find, especially for a small museum. In addition, HTML coding is susceptible to small errors that can cause big headaches. A misplaced bracket or a missing backslash can cause hours of lost time spent searching through lines of code. While there are HTML-authoring software solutions like DreamWeaver that offer so-called WYSIWYG (What You See Is What You Get) modes, they offer only a very limited

range of options when used in this mode. You can build a very simple HTML-based site using a WYSIWYG editor, but it still takes a basic understanding of what HTML code can provide in order to create an effective site.

Because of its steep learning curve, depth of required expertise, and time-consuming nature, I rejected HTML as an option for the online exhibit.

Option two: Virtual Exhibit add-on for PastPerfect

The next option was to install Virtual Exhibit, an add-on available for PastPerfect Museum Software. According to the description on the PastPerfect Web site, the Virtual Exhibit add-on utilizes existing PastPerfect records to create HTML code that can be used to create an online collection. The site also promises that no experience in HTML or Web programming is necessary, and that users can have a customized online collection available in “just a few minutes” (www.museumsoftware.com/ve4.shtml).

The Railroad and Heritage Museum has only recently begun using PastPerfect to catalogue and track its collections, and the Houston Public Library collection of maps had not been entered into the system at the time this collaboration began. Therefore, there were no existing records in the system that could have been used to create an online collection using the Virtual Exhibit add-on. A great deal of time would therefore have to be invested in creating PastPerfect records for each map before moving forward with the online exhibit, a process that could only take place on-site at the RHM. In addition, the cost to add Virtual Exhibit is \$420, a significant expense for a small institution, especially one without prior experience creating a collections-based online presence.

Due to its fairly substantial cost and the requirement of having to enter records on-site at the RHM, I rejected this option.

Option three: Pachyderm

Pachyderm is a template-based software that uses Adobe Flash to create online exhibits. It is designed specifically for museums and other collecting institutions, especially those whose staffs have little or no multimedia programming expertise. Billed as a solution “as easy to use as filling out a web form,” Pachyderm “began as a partnership led by The New Media Consortium (NMC) and the San Francisco Museum of Modern Art (SFMOMA), and funded by the Institute for Museum and Library Services (IMLS)” (<http://pachyderm.nmc.org/about/>). The Pachyderm software is open source, but an annual fee of \$99 to \$149 applies if a museum wants to host its Pachyderm presentation on the NMC servers.

I became aware of Pachyderm through discussions with other Museum Studies students, and was pleasantly surprised to learn that Dr. Gardner Campbell, Director of Baylor’s Academy for Teaching and Learning, was also a member of the New Media Consortium’s board. Dr. Campbell graciously offered to serve as a contact with the staff at NMC who administer Pachyderm, and after an informative conference call with Dr. Campbell, myself, NMC staff member Rachel K. Varon, and colleagues from the Martin Museum of Art, I saw the advantages of using Pachyderm to author my online exhibit.

However, the \$99 price tag was an obstacle, as I had hoped to find a free or very low-cost alternative for this part of the project. Dr. Campbell’s inquiries to his colleagues at NMC revealed that, because Baylor University is a full member of the consortium, the university is entitled to five free sets of Pachyderm credentials. This fortuitous circumstance eliminated the economic barrier hitherto described, and I was able to choose Pachyderm as my online exhibit authoring solution with no hesitation.

Utilizing Pachyderm

I received my login credentials on January 26, 2010 and began work on the online exhibit the next day. As was the case with Flickr, everything involved with the Pachyderm process is interfaced through a Web site. Each Pachyderm account – in this case, a Baylor University account assigned to me – receives hosting on NMC servers, which means there is no need for museum staff to install, configure, or monitor local servers to host their exhibit. The \$99 annual fee helps NMC cover these costs.

Pachyderm allows users to choose from a number of templates to create an exhibit that best displays the digitized resources. Once a homepage template (See Figure 7) is selected, additional dependent pages are created that link back to it, making for a clean, linear progression from page to page. For multimedia elements like photos and videos, users upload files to the NMC servers through a simple interface, add metadata to the elements to make them findable, and then plug them into the templates where appropriate. This off-site storage means users do not have to invest time and energy making sure source files are always available; the NMC servers handle this remotely.

Before I began entering data into the templates, I first created a Word document where I sketched out how I wanted the online exhibit to be organized. Using this document as a guide, I created a homepage titled “Railroad and Heritage Museum: Vintage Survey Maps” and six dependent pages:

1. About the Collection: Describes how the maps came to the RHM, how they are used, and one sample image
2. How Were They Used?: Describes ways the railroads used the maps, such as for identification of landowners and to show locations of depots and stations

3. Railroads in Texas, 1949: A “zoom” template featuring a 1949 map showing major rail lines in the state of Texas
4. Take A Closer Look: A “zoom” template featuring a c. 1920 map of Houston and its many rail yards
5. Different Draftsmen, Different Styles: A discussion of the aesthetics of the maps, as well as seven different sample maps for closer examination
6. Want to Learn More?: Features a link to the Flickr DAMS site, a page describing why the maps were digitized, a page showcasing the Cruse large format scanner, and a link to the Railroad and Heritage Museum’s Web site (See Figure 8)

The section titles contain a mix of declarative statements (Different Draftsmen, Different Styles) and questions (Want to Learn More?) in order to grab viewer interest and to avoid a series of simple declarative statements.

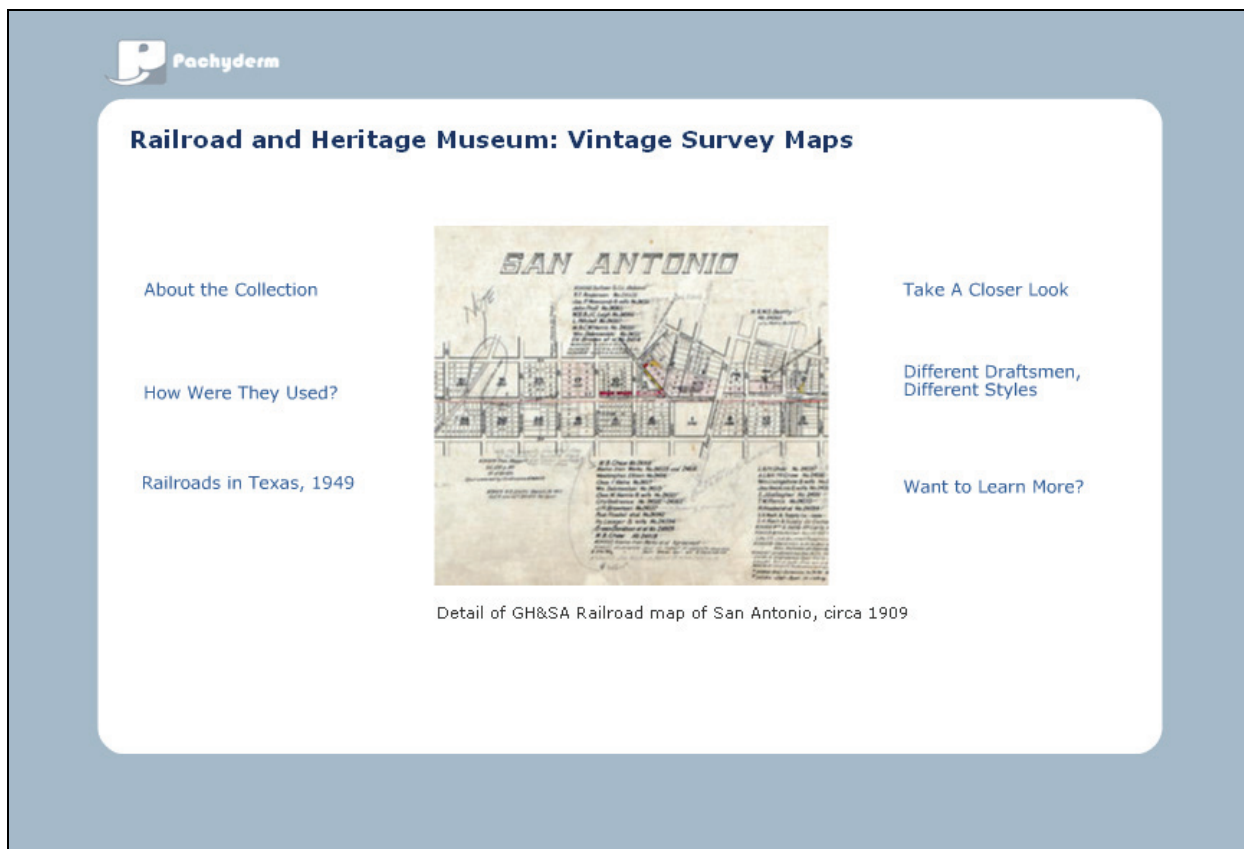


Figure 7: The online exhibit homepage, via Pachyderm

Once the pages were created, I began re-sizing images to fit the relatively small file sizes allowed by the Pachyderm software. For example, images (which must be in .gif or .jpg format) are only allowed to reach 2 MB per file, with a maximum display size of 2560 x 2648 pixels. Once the images were resized, I uploaded them to the “Media” section of the software so they could be plugged into the templates. As each dependent page was created, I linked it back to the homepage. In this manner, I was able to finish the entire process – including image resizing and upload; text creation; template formatting; and final rendering – in approximately six hours. The process was sufficiently well explained by both on-screen prompts and helpful PDF files available on the Pachyderm Web site that I feel confident users with even minimal computer proficiency could create a similar site in 12 hours or less. In this way, Pachyderm definitely lived up to its billing as the simplest way for museums to create online exhibits.

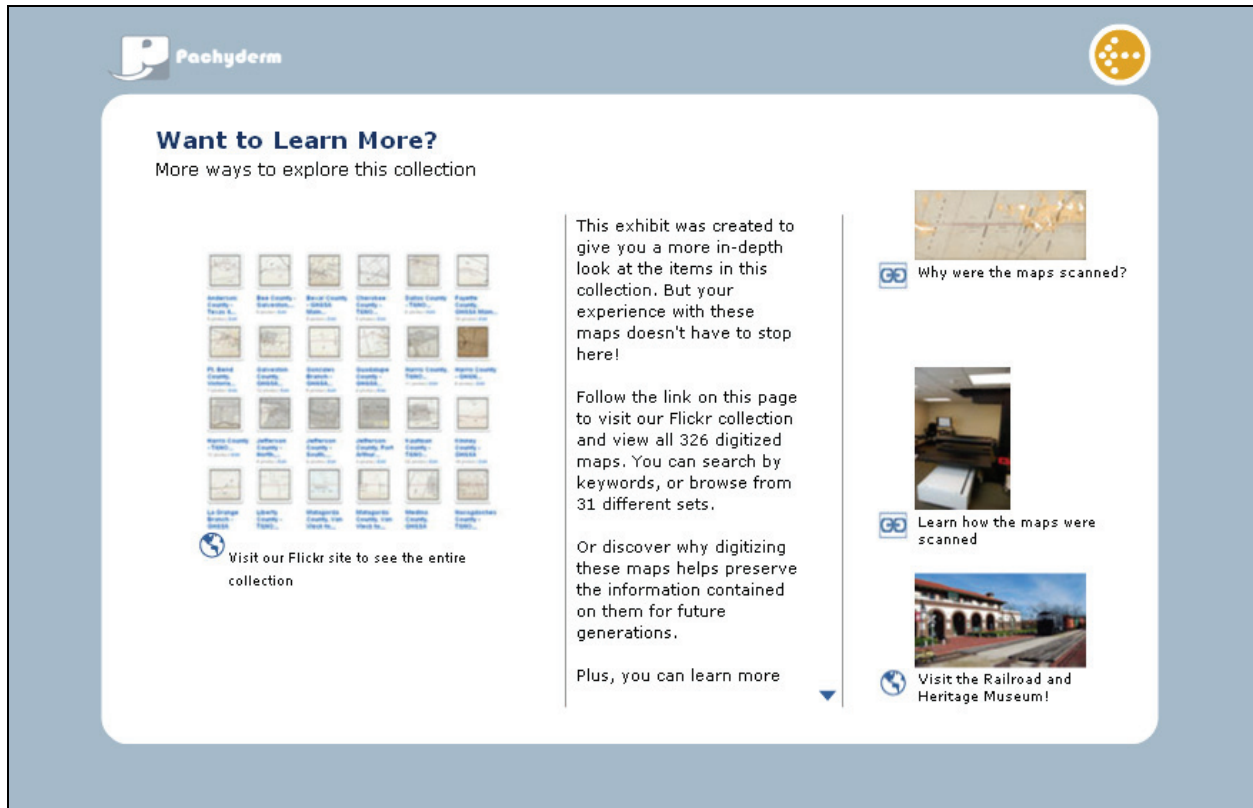


Figure 8: The “Want to Learn More?” slide from the online exhibit

The completed presentation lives on the NMC server at services.pachyderm.org,¹⁶ and it is a simple matter of sending along the URL to the Webmaster at the Railroad and Heritage Museum to get it onto their homepage. As with the decision to use Flickr, the use of Pachyderm minimizes risk to the RHM, presents a fairly robust solution to the issue presented (i.e. the creation of an online exhibit without complex coding), and is cost effective.

¹⁶ The Pachyderm site is available from: <http://services.pachyderm.org/pachyderm/presentations/00-4867-29-12003165-6000013810336-13363665-57-74/>

CHAPTER FIVE:

Conclusions

This project included a number of elements, including the selection of an appropriate museum partner, the digitization of hundreds of unique resources using current best practices, and the placement of those resources in two different online systems for access by a worldwide audience. Over the course of this process, I have arrived at some conclusions that may impact how small to medium-sized museums move forward with similar projects.

I believe that museums of any size can – and should – benefit from digitizing their collections, especially rare or fragile archival materials. The benefits, including preserving information, providing backup copies for display or reproductions, and providing greater access via the Internet, are worth the investments in time, money, and staff training.

Because desktop scanners have become widespread and affordable in recent years, museums of almost any size should be able to afford to purchase one. By making allowances in the annual budget ahead of time, any institution can make plans to add a moderately powerful digitization tool to their collections area, as part of the work responsibilities for the registrar, collections manager, or curator. The basic “plug and play” nature of most desktop scanners allows for their use with most personal computers, and many include software that allows users to perform basic edits such as cropping, levels adjustments, and saving in different formats. These capabilities should be more than ample for most museums to begin digitizing their archival collections.

However, in the case of oversized, extremely fragile, or other challenging materials, it often is not possible for a small to medium-sized museum to purchase the specialized scanners required to digitize them. For example, the Cruse large format scanner used in this project comes with a hefty price tag: \$180,000. Add in the time required to ship (3-6 weeks) and install (3 days) the unit, plus training time (1 full working day), and the investment is far beyond the reach of even some larger museums. While the Cruse is an extreme example, most other specialized scanners come with a high price tag, complex training, or long lead times for installation, if not all three.

Because of the challenges listed above, collaboration and partnerships are much more likely to bear fruit for small museums. In the case of this project, my partnership with the Railroad and Heritage Museum gave them access to the resources of the Riley Digitization Center at Baylor University, a center equipped with specialized scanners, highly-trained staff, and the luxury of a healthy annual operating budget. While the RDC currently focuses its efforts on digitizing materials from the large holdings of Baylor's special collections, it is likely that this project will serve as a model for future outreach to local organizations, much along the lines of the work being doing by the University of North Texas' Digital Projects Unit.¹⁷

I believe that larger institutions such as Baylor University, the University of North Texas (UNT), the University of Texas at Austin, Texas Tech University and others that have made significant investments in digitization technology will be seeking to form mutually beneficial partnerships with greater frequency in the coming years. An agreement wherein the digitizing entity performs the work in exchange for adding the digitized materials to its own collections (while providing access copies and

¹⁷ The UNT Digital Projects Unit. University of North Texas. Denton, TX. Accessed October 8, 2009. Available from <http://www.library.unt.edu/digitalprojects>.

proper citations to the lending museum) seems to be the most likely scenario, although a fee-based digitization service like the one at UNT is also a possibility.

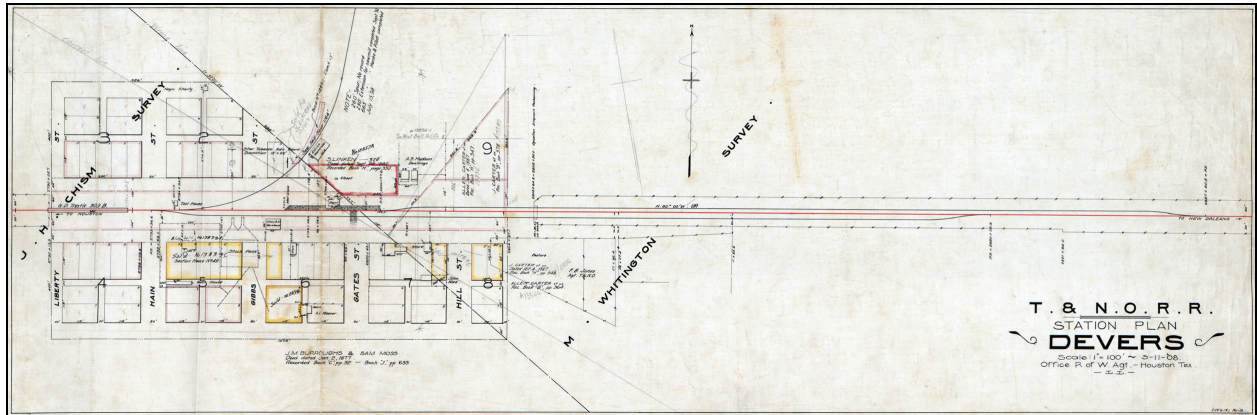
A note concerning costs. To amend one of Mayborn Museum director Ellie Caston's favorite phrases, I am fully aware that the Baylor University Electronic Library "is not the real world." It has the benefit of endowed funds, an active fundraising effort, and an established annual budget to support its efforts. Therefore, many of the things the DPG is able to do for its clients are beyond the reach of most digitization efforts. But I also disagree with the notion that only large institutions are able to afford such luxuries. If a museum is dedicated to carrying out digitization projects, it will find the funds it needs to achieve its goals, just as it does when a particularly appealing traveling exhibit is available, or a new wing is required to house a major new acquisition. One method may be through grants, both governmental and private. Many grant funders have begun to require an element of digitization or online access as part of its granting process, and digitization-focused projects are often given priority because of their proven effectiveness at connecting users with resources.

Ultimately, I believe the outcomes realized from a proper digitization project are well worth the time, money, and training invested to carry them out. If a museum is, at its most basic level, a steward of our cultural heritage, then the greatest part of that stewardship must involve preserving and making available those resources which it has been entrusted to hold. Digitization offers a proven, effective way of achieving both of these ends.

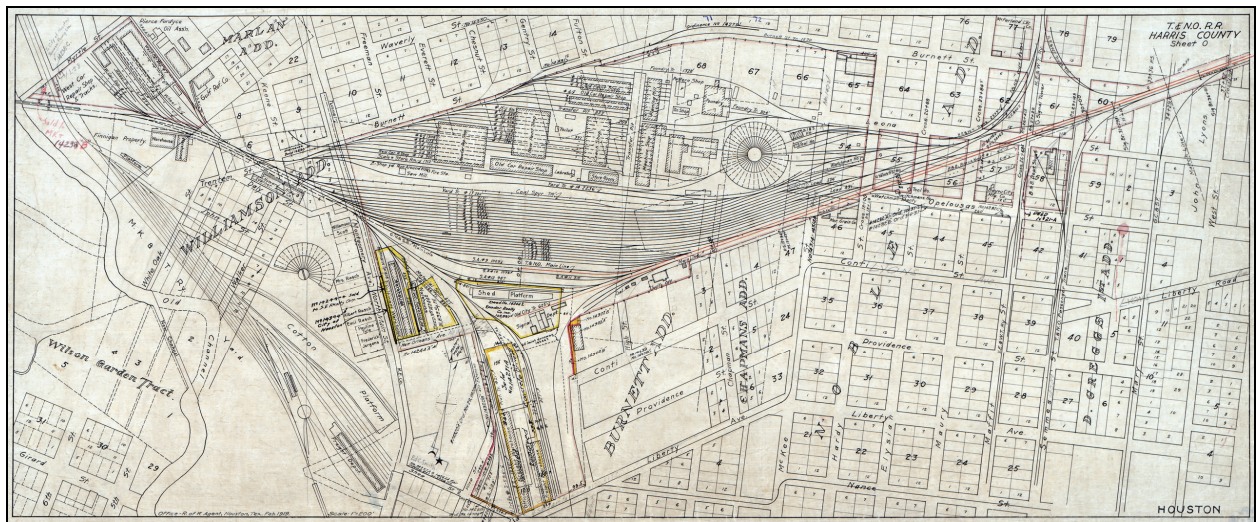
In summary, I believe that even the smallest museums have the opportunity and the responsibility to digitize their archival collections, whether through investing in their own equipment and training or through partnership with a larger entity. The security and usability of our cultural heritage demands nothing less.

APPENDIX A

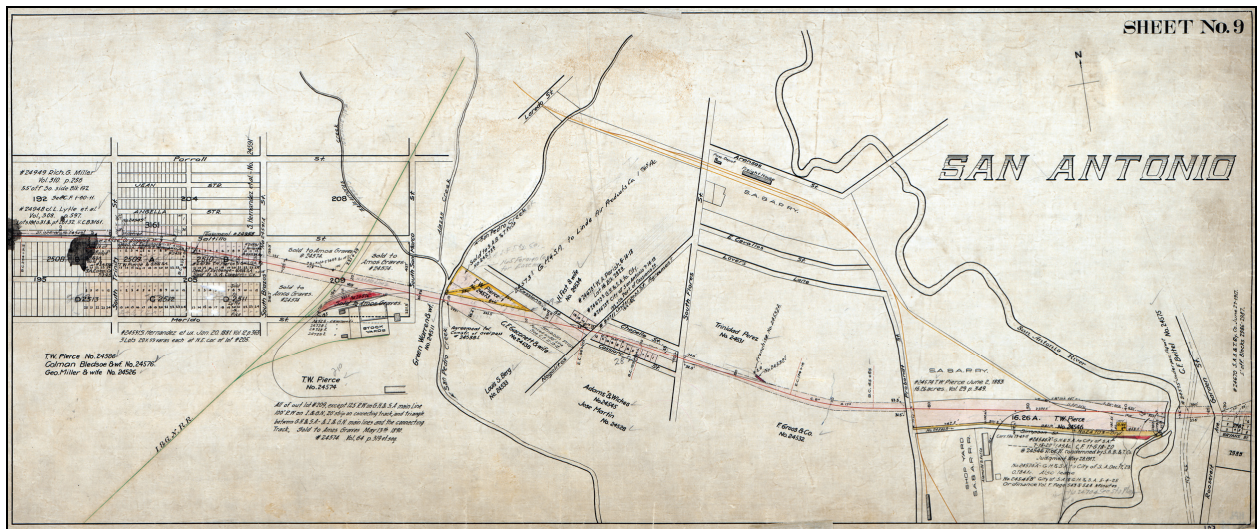
The RHM Houston Public Library Collection: Selections from the Collection



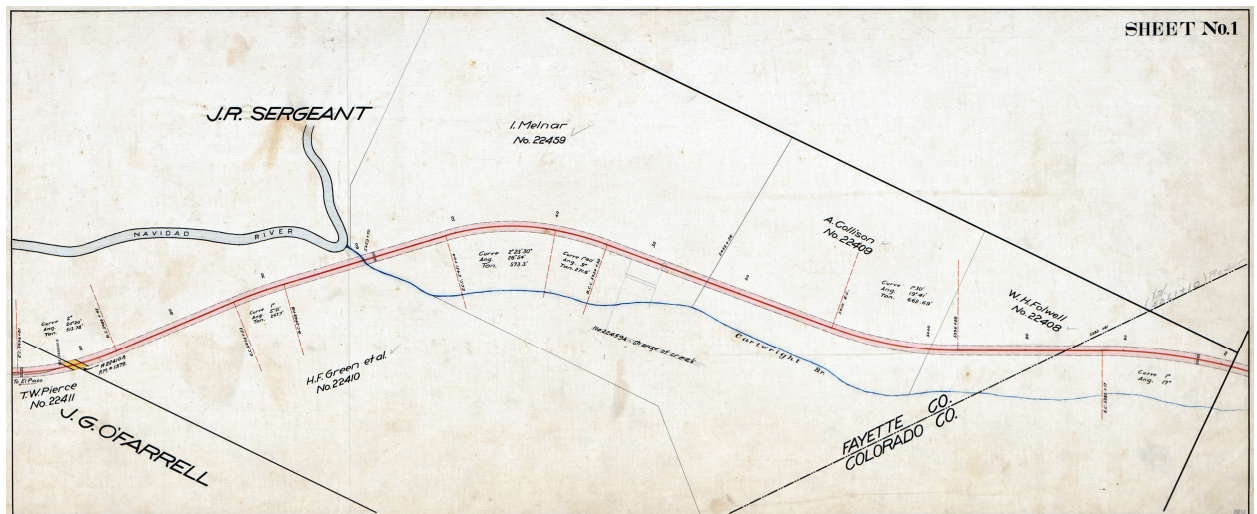
T&NO Railroad station plan of Devers, Texas. Project ID# rhm_hpl_38_1006. Drafted May 5, 1908.



T&NO Railroad map of sections of Houston, Texas. Project ID# rhm_hpl_25_781.



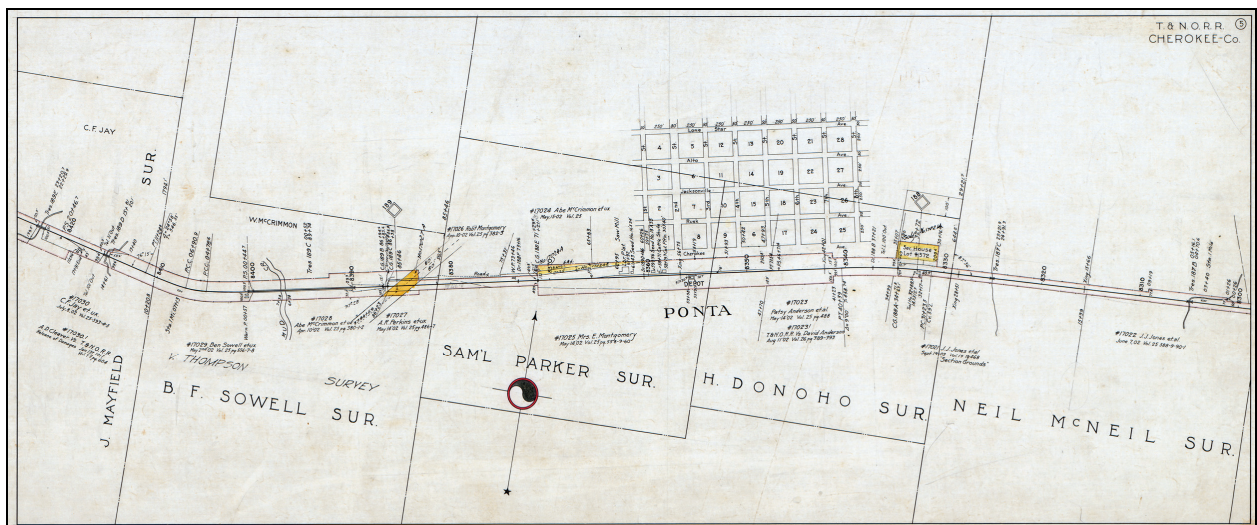
GH&SA Railroad map of portions of San Antonio, Texas. Project ID# rhm_hpl_05_521.



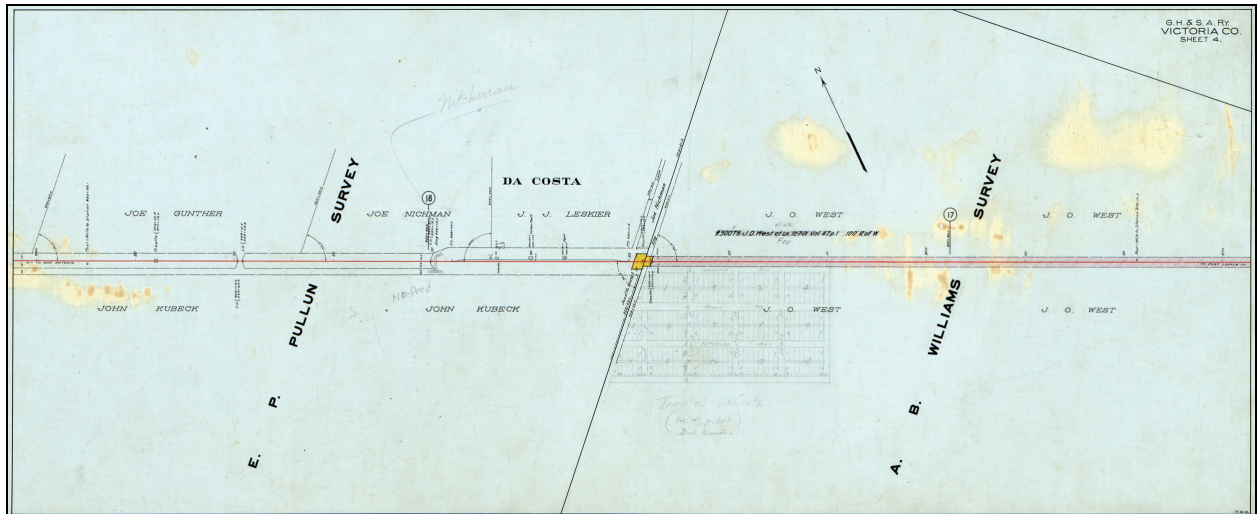
GH&SA Railroad map of portions of Fayette and Colorado counties, Texas. Project ID# rhm_hpl_15_670.



GH&SA Railroad map of portions of Galveston County, Texas. Project ID# rhm_hpl_19_717.



T&NO Railroad map of portions of Cherokee County, Texas. Project ID# rhm_hpl_085_545.



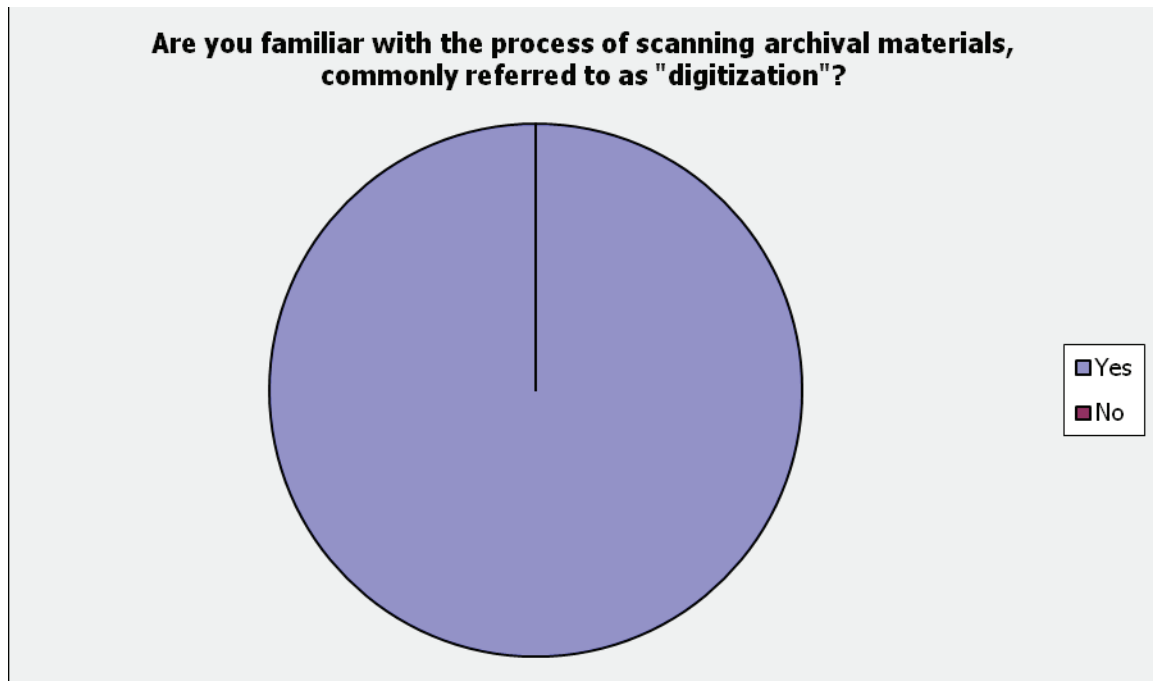
GH&SA Railroad map of portions of Victoria County, Texas. Project ID# hpl_rhm_49_1185.

APPENDIX B

Survey information

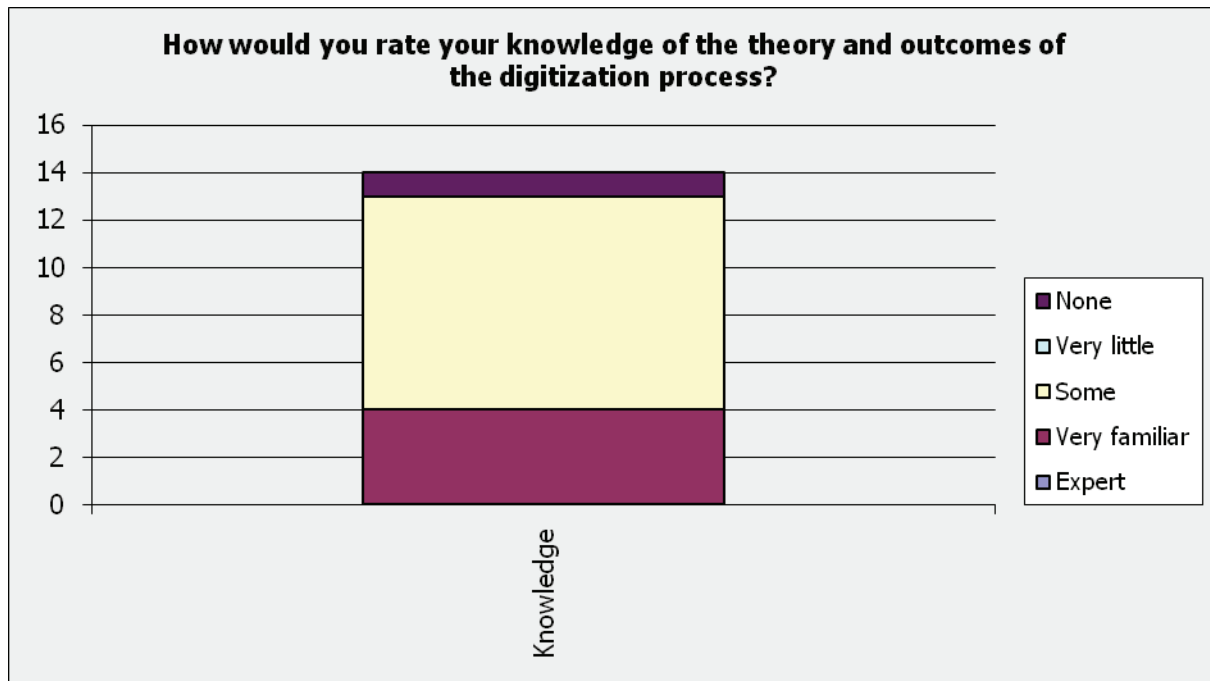
Question 1

Are you familiar with the process of scanning archival materials, commonly referred to as "digitization"?		
Answer Options	Response Percent	Response Count
Yes	100.0%	14
No	0.0%	0
<i>answered question</i>		14
<i>skipped question</i>		0



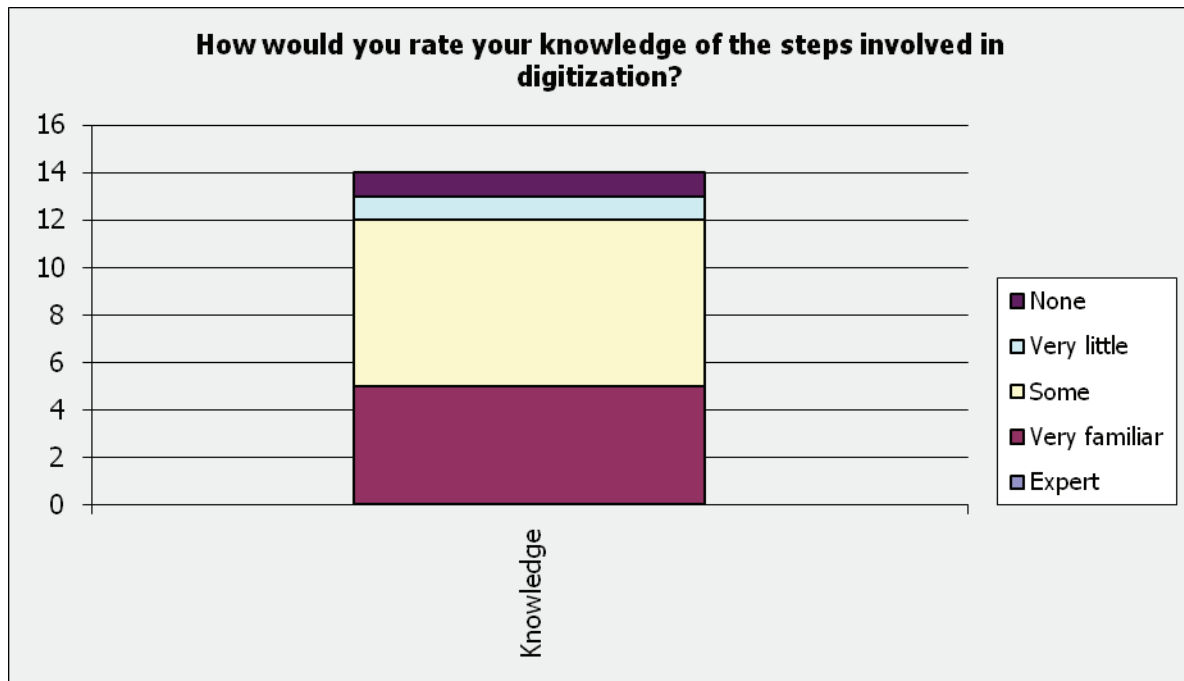
Question 2

How would you rate your knowledge of the theory and outcomes of the digitization process?						
Answer Options	None	Very little	Some	Very familiar	Expert	Response Count
Knowledge	1	0	9	4	0	14
<i>answered question</i>						14
<i>skipped question</i>						0



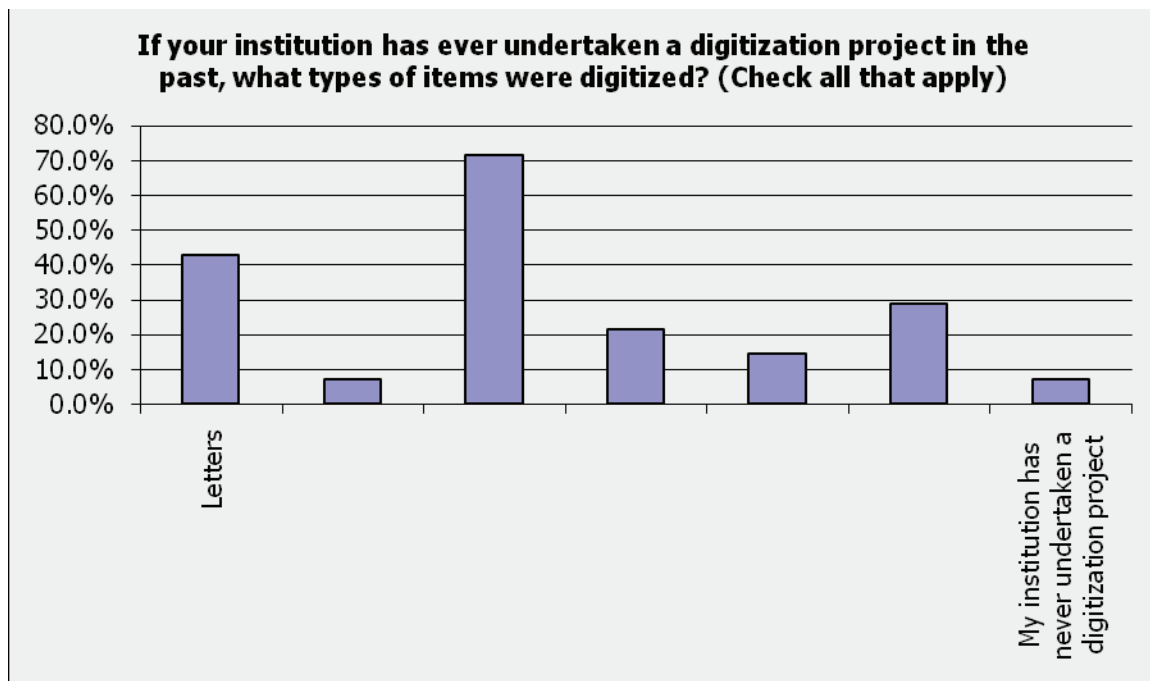
Question 3

How would you rate your knowledge of the steps involved in digitization?						
Answer Options	None	Very little	Some	Very familiar	Expert	Response Count
Knowledge	1	1	7	5	0	14
<i>answered question</i>						14
<i>skipped question</i>						0



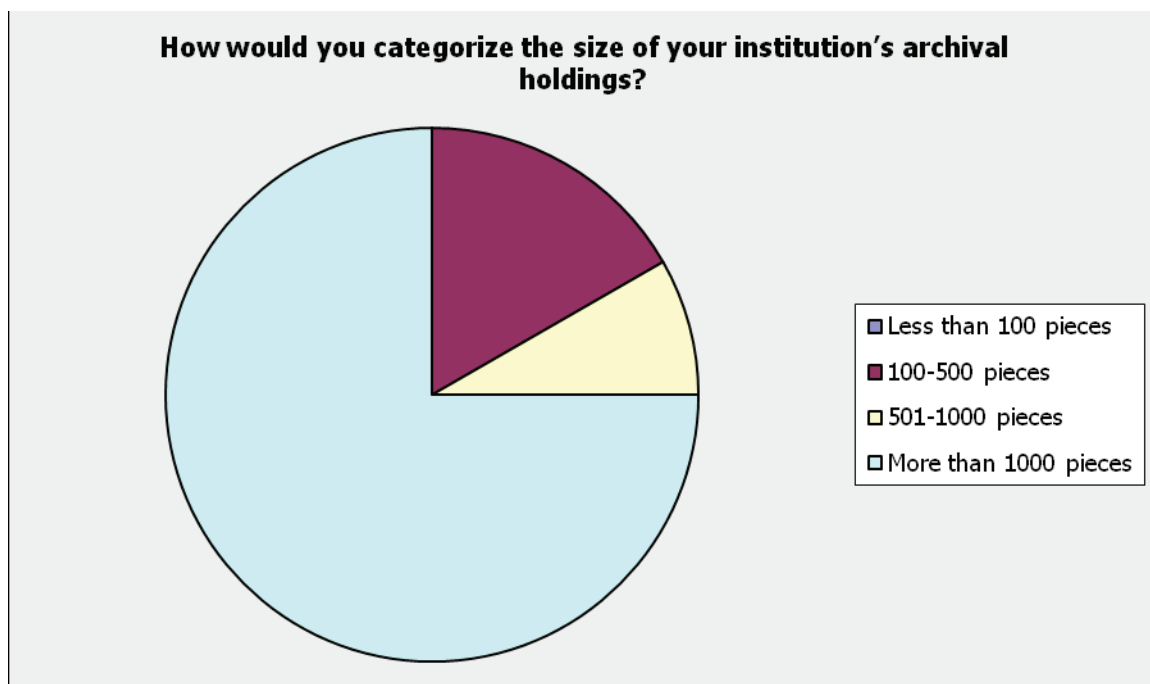
Question 4

If your institution has ever undertaken a digitization project in the past, what types of items were digitized? (Check all that apply)		
Answer Options	Response Percent	Response Count
Letters	42.9%	6
Maps	7.1%	1
Photographs	71.4%	10
Books	21.4%	3
Government documents	14.3%	2
3D objects	28.6%	4
My institution has never undertaken a digitization project	7.1%	1
Other (please specify)		3
answered question		14
skipped question		0



Question 5

How would you categorize the size of your institution's archival holdings?		
Answer Options	Response Percent	Response Count
Less than 100 pieces	0.0%	0
100-500 pieces	16.7%	2
501-1000 pieces	8.3%	1
More than 1000 pieces	75.0%	9
answered question		12
skipped question		2



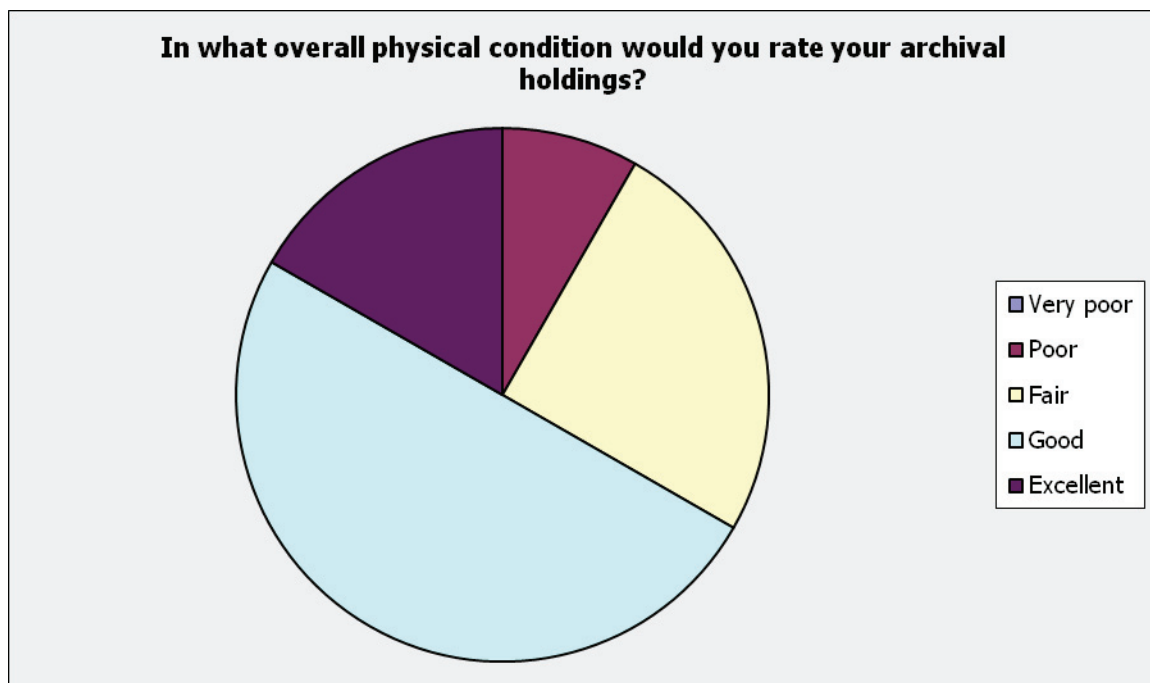
Question 6

What types of objects are included in your archival holdings? (Check all that apply):		
Answer Options	Response Percent	Response Count
Letters	75.0%	9
Maps	66.7%	8
Photographs (loose)	100.0%	12
Photo albums (bound)	58.3%	7
Official documents (deeds, letters of commission, etc.)	75.0%	9
Diaries, journals, account books, etc.	66.7%	8
Other books	75.0%	9
Scrapbooks	66.7%	8
Pamphlets, brochures, promotional materials	83.3%	10
Other non-3D objects (commemorative ribbons, packaging, labels, etc.)	41.7%	5
Other (please specify)		1
answered question		12
skipped question		2



Question 7

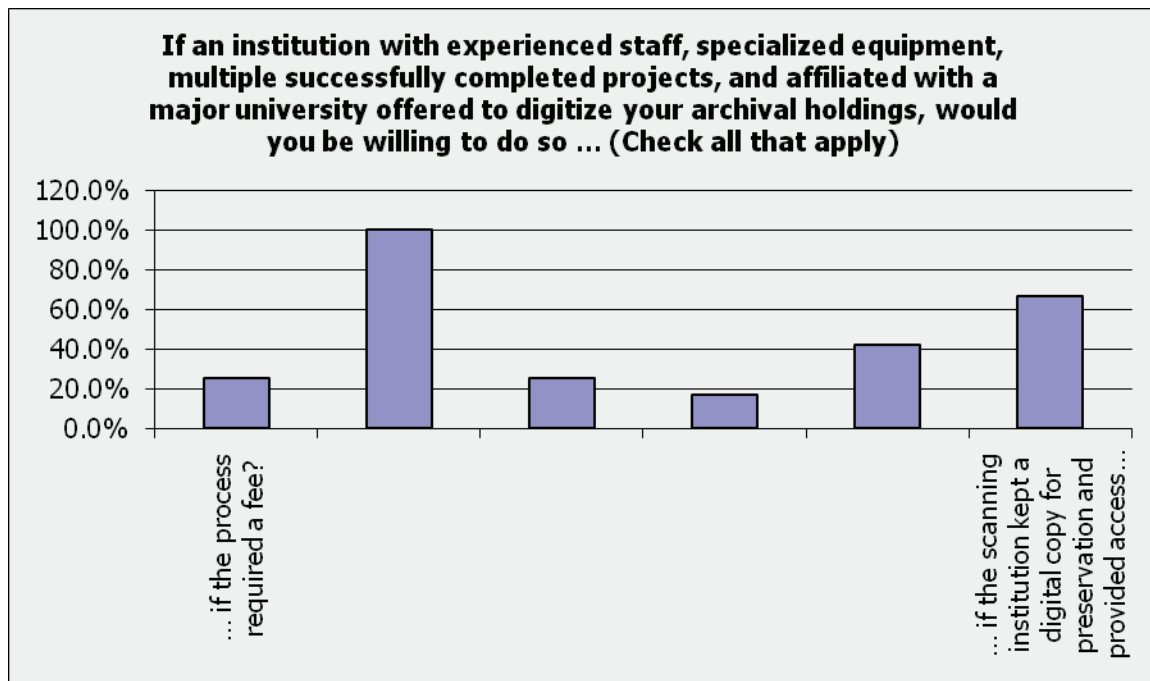
In what overall physical condition would you rate your archival holdings?		
Answer Options	Response Percent	Response Count
Very poor	0.0%	0
Poor	8.3%	1
Fair	25.0%	3
Good	50.0%	6
Excellent	16.7%	2
<i>answered question</i>		12
<i>skipped question</i>		2



Question 8

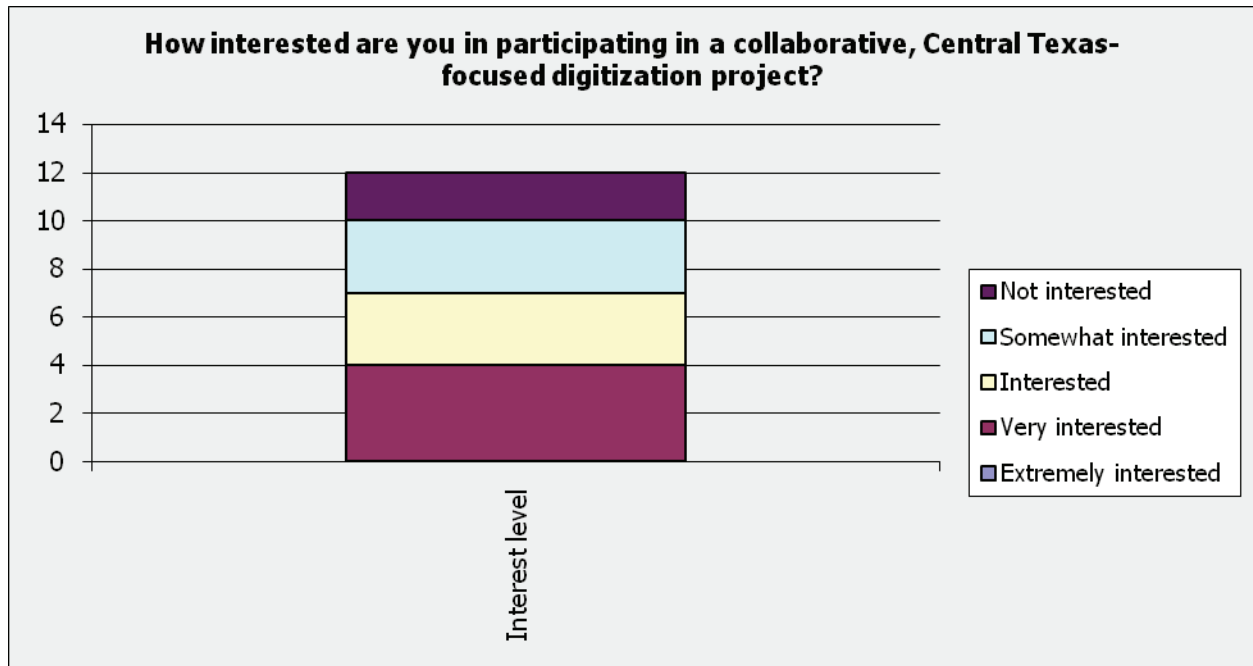
If an institution with experienced staff, specialized equipment, multiple successfully completed projects, and affiliated with a major university offered to digitize your archival holdings, would you be willing to do so ... (Check all that apply)

Answer Options	Response Percent	Response Count
... if the process required a fee?	25.0%	3
... if the process were free?	100.0%	12
... if the process required sending your archival material no more than 100 miles away from your institution?	25.0%	3
... if the scanning institution wanted to host digital copies of your material (credited to your institution) on its digital collections Web site?	16.7%	2
... if the scanning institution kept a digital copy of the scanned items but did NOT place them online (i.e., only for preservation, not access)?	41.7%	5
... if the scanning institution kept a digital copy for preservation and provided access copies to your institution for hosting on your own Web site?	66.7%	8
Other (please specify)		1
answered question		12
skipped question		2



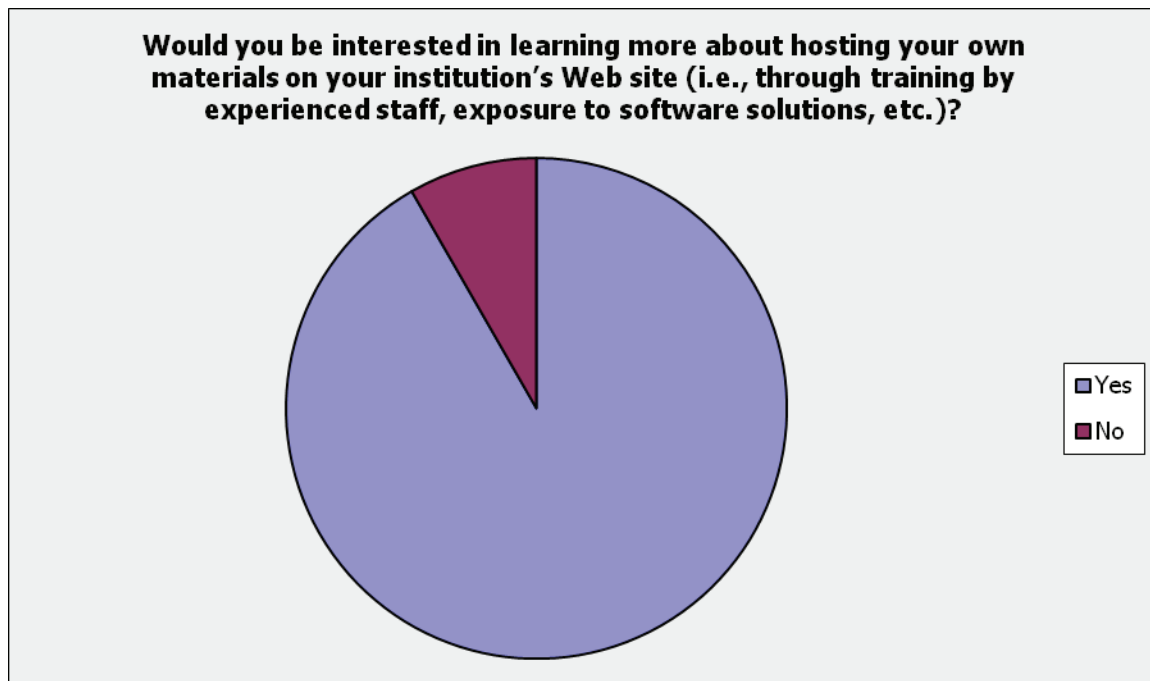
Question 9

How interested are you in participating in a collaborative, Central Texas-focused digitization project?						
Answer Options	Not interested	Somewhat interested	Interested	Very interested	Extremely interested	Response Count
Interest level	2	3	3	4	0	12
<i>answered question</i>						12
<i>skipped question</i>						2



Question 10

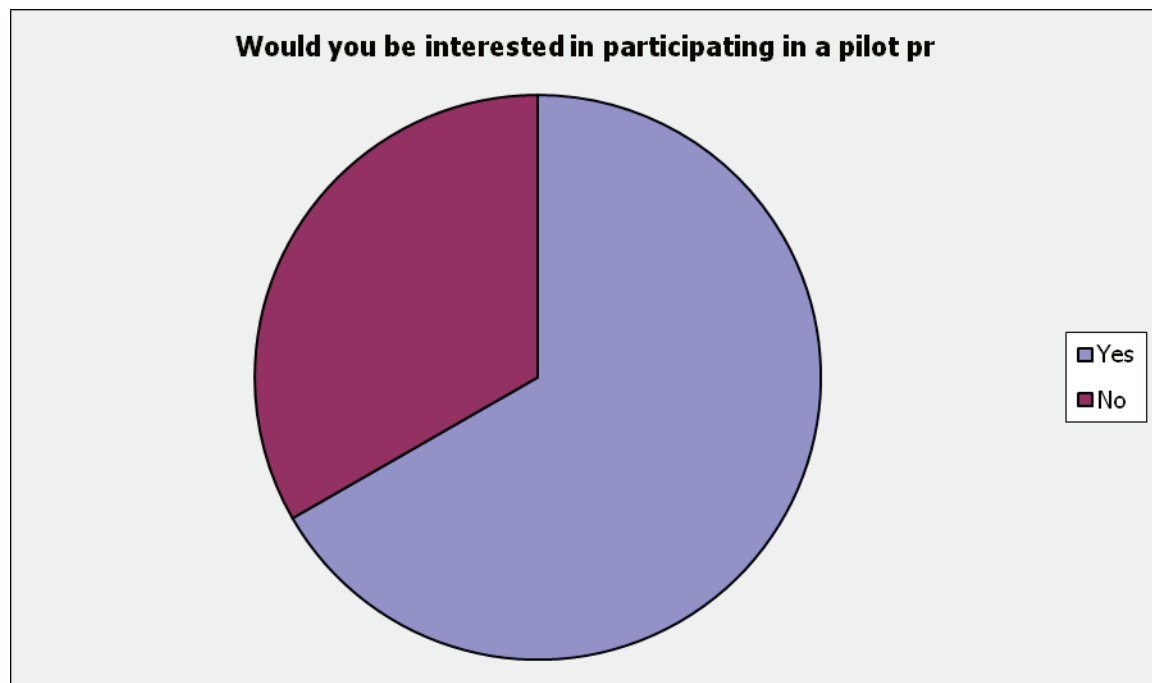
Would you be interested in learning more about hosting your own materials on your institution's Web site (i.e., through training by experienced staff, exposure to software solutions, etc.)?		
Answer Options	Response Percent	Response Count
Yes	91.7%	11
No	8.3%	1
<i>answered question</i>		12
<i>skipped question</i>		2



Question 11

Would you be interested in participating in a pilot project that would serve as a model for future collaborative digitization projects in Central Texas? (Note: a “yes” response does not guarantee participation; further planning and consultation with the author of this survey would be required to participate.)

Answer Options	Response Percent	Response Count
Yes	66.7%	8
No	33.3%	4
<i>answered question</i>		12
<i>skipped question</i>		2



Question 12

How many full-time employees are on-staff at your institution? Part-time? How many volunteers?		
Answer Options	Response Percent	Response Count
Full-time	90.0%	9
Part-time	70.0%	7
Volunteers	90.0%	9
<i>answered question</i>		10
<i>skipped question</i>		4



Average number of full-time staff: 6.3

Average number of part time staff: 4.4

Average number of volunteers: 21.2

Question 13

On average, what is your institution's total annual attendance? (Estimates are fine.)	
Answer Options	Response Count
	8
<i>answered question</i>	8
<i>skipped question</i>	6

Responses:

65,000

55,000

15,000

18,000

35,000

5,000-6,000

23,000

1,500

Question 14

What year was your institution founded?	
Answer Options	Response Count
	10
<i>answered question</i>	10
<i>skipped question</i>	4

Responses:

1991

1992

1935

1997

1991

1978

Before 1971

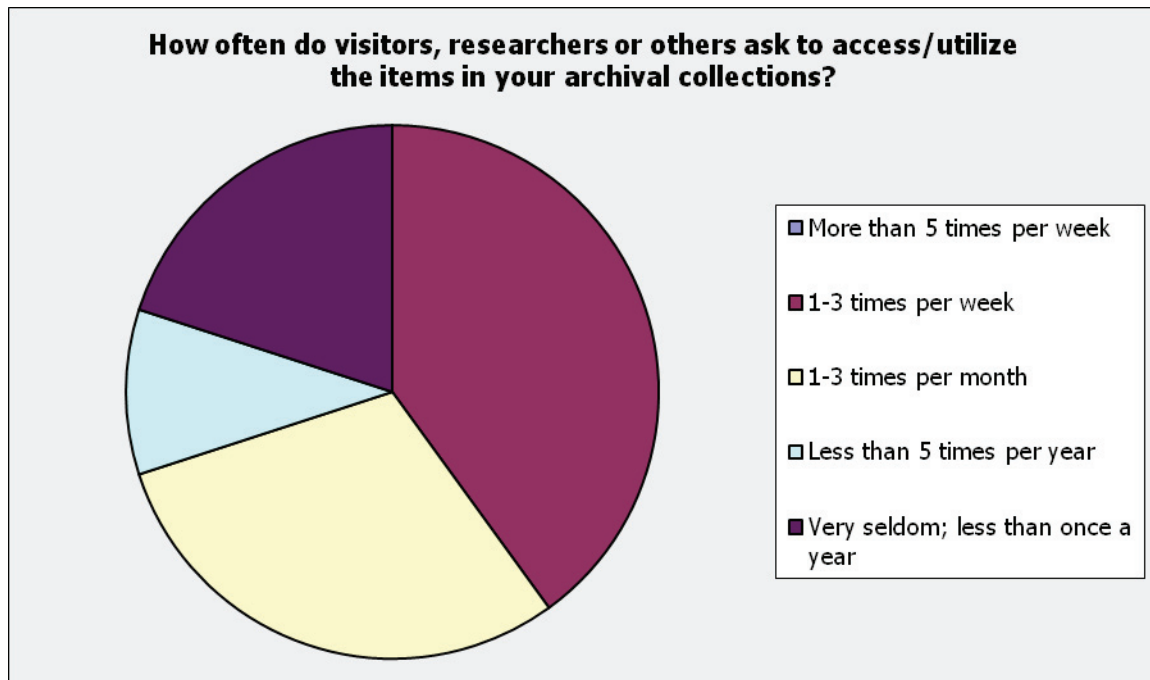
1968

1989

1971

Question 15

How often do visitors, researchers or others ask to access/utilize the items in your archival collections?		
Answer Options	Response Percent	Response Count
More than 5 times per week	0.0%	0
1-3 times per week	40.0%	4
1-3 times per month	30.0%	3
Less than 5 times per year	10.0%	1
Very seldom; less than once a year	20.0%	2
answered question		10
skipped question		4



Question 16

Does your institution have a Web site? If so, do you currently host any online content from your institution's collections – virtual exhibits, links to scans of archival material, etc.?		
Answer Options	Response Percent	Response Count
My institution does not maintain a Web site	0.0%	0
My institution has a Web site but does not host any online content	70.0%	7
My institution has a Web site and currently hosts online content	30.0%	3
<i>answered question</i>		10
<i>skipped question</i>		4

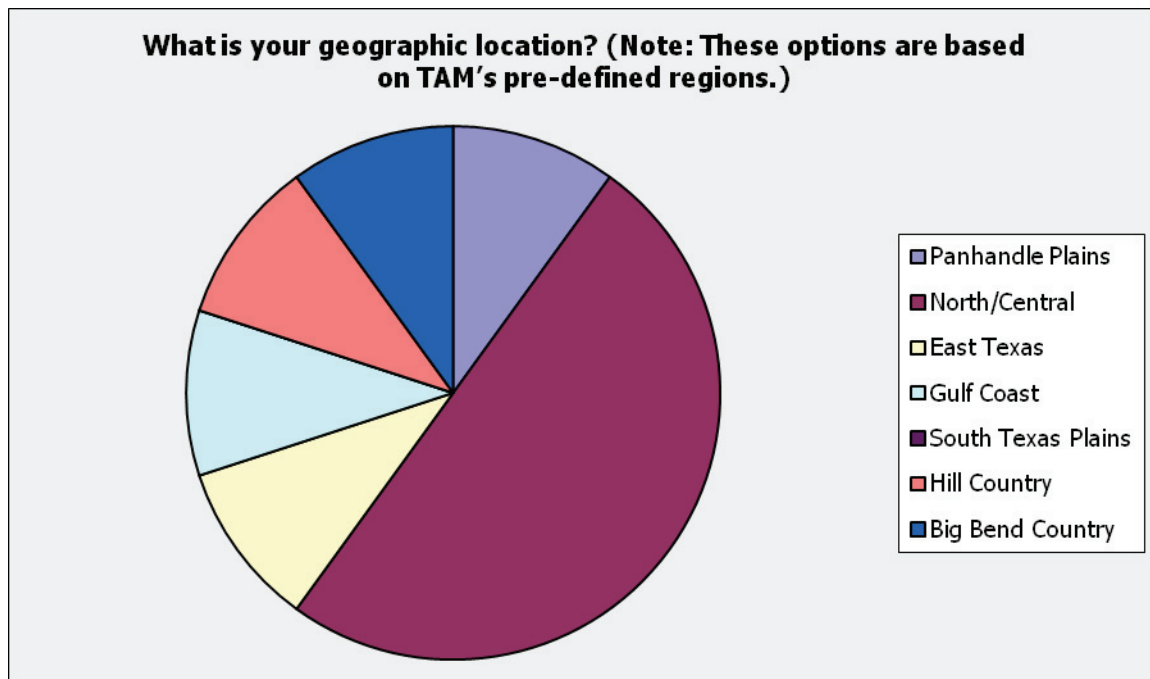
Question 17

Are you – or any of your staff members – associated with the Texas Association of Museums, the Society of Southwest Archivists, or a similar professional organization? If so, please specify.		
Answer Options	Response Percent	Response Count
No one on staff is associated with these organizations	0.0%	0
Texas Association of Museums	100.0%	10
Society of Southwest Archivists	0.0%	0
Other (please specify)		1
<i>answered question</i>		10
<i>skipped question</i>		4

Other: American Association of Museums

Question 18

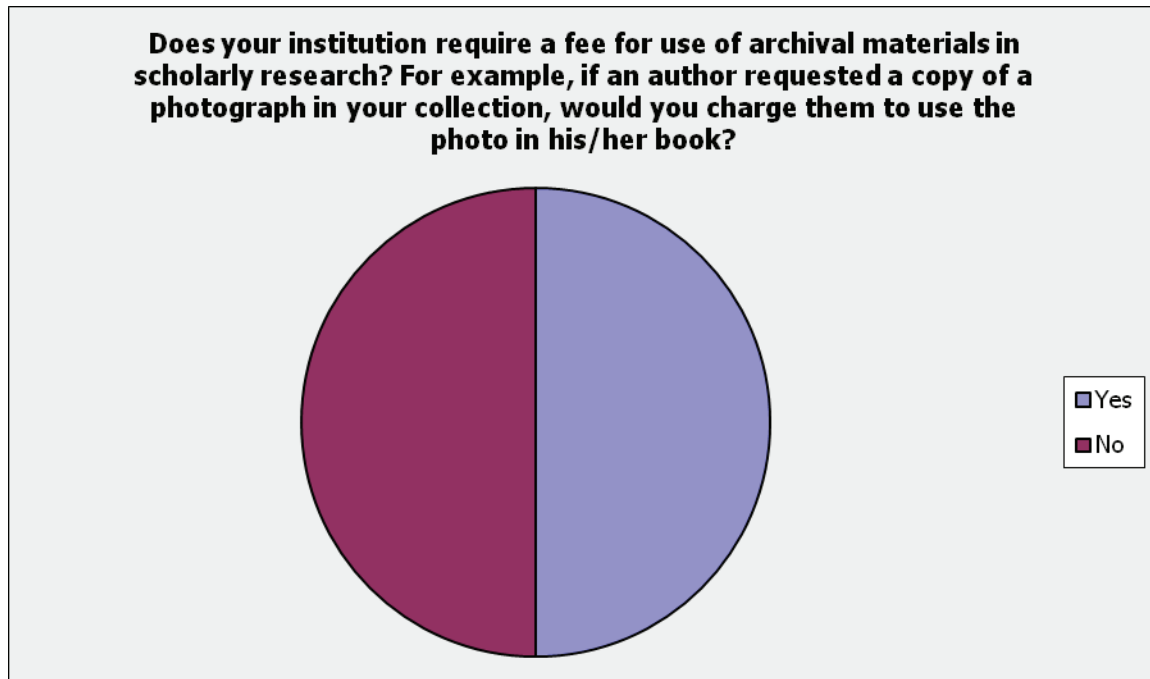
What is your geographic location? (Note: These options are based on TAM's pre-defined regions.)		
Answer Options	Response Percent	Response Count
Panhandle Plains	10.0%	1
North/Central	50.0%	5
East Texas	10.0%	1
Gulf Coast	10.0%	1
South Texas Plains	0.0%	0
Hill Country	10.0%	1
Big Bend Country	10.0%	1
answered question		10
skipped question		4



Question 19

Does your institution require a fee for use of archival materials in scholarly research? For example, if an author requested a copy of a photograph in your collection, would you charge them to use the photo in his/her book?

Answer Options	Response Percent	Response Count
Yes	50.0%	5
No	50.0%	5
<i>answered question</i>		10
<i>skipped question</i>		4




APPENDIX C

Pachyderm Site Screen Captures

Note: Live site available online at <http://services.pachyderm.org/pachyderm/presentations/00-4867-29-12003165-6000013810336-13363665-57-74/>

Home page

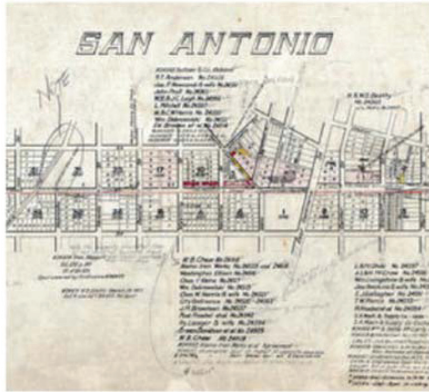


Railroad and Heritage Museum: Vintage Survey Maps

About the Collection

How Were They Used?

Railroads in Texas, 1949



Take A Closer Look

Different Draftsmen,
Different Styles

Want to Learn More?

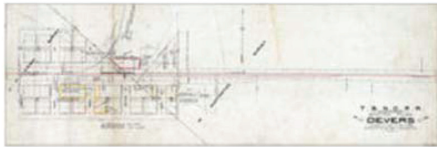
Detail of GH&SA Railroad map of San Antonio, circa 1909

About the Collection



About the Collection

The RHM collection of early 20th century railroad survey maps



Texas & New Orleans Railroad plan of Devers, Texas. Drafted May 11, 1908.

The maps in our collection were donated to the Railroad Heritage Museum by George Werner, an avid collector of railroad-related materials. He acquired them from the office of the Southern Pacific Railroad when the SPR was preparing to dispose of them. Werner then gave the maps -- known as the Houston Public Library collection -- to the RHM.

Prized by researchers and "railfans" alike, these maps show rights-of-way for several historically important Texas railroads. Major Texas towns such as Houston, Dallas, and San Antonio are represented, as well as town sites with names like Sacul, Zillah, Raisin and Don



How Were They Used?



How Were They Used?



Detail of map of Dallas, Texas, dated March 16, 1909.

The maps were used by railroad companies as a reference to engineering teams, surveyors, and planners whose decisions guided the growth and expansion of the lines.

For example, the maps indicate the names of landowners along the rights-of-way of each line. This information would help the company negotiate the rights to build new lines across new sections of land.

Because they list physical features like lakes, rivers, and elevation changes, the maps would aid construction crews and surveyors as they plotted the tracks' route across the state.



Identify natural features like lakes and rivers

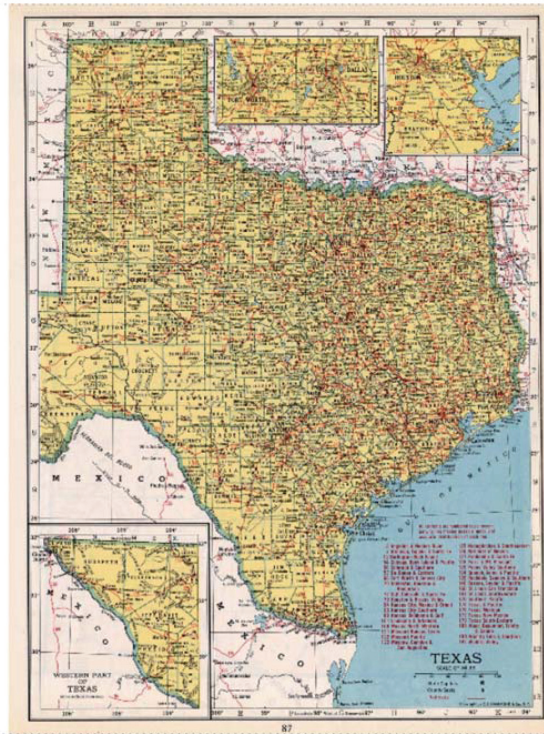


Show location of depots and stations

A table listing deed holders in Victoria, Texas. The table has multiple columns with headers including "DEED", "REFERENCE", "DATE", "NAME", "ADDRESS", "CITY", "COUNTY", "STATE", and "REMARKS". The table contains several rows of data.

List of deed holders, Victoria, TX

Railroads In Texas, 1949

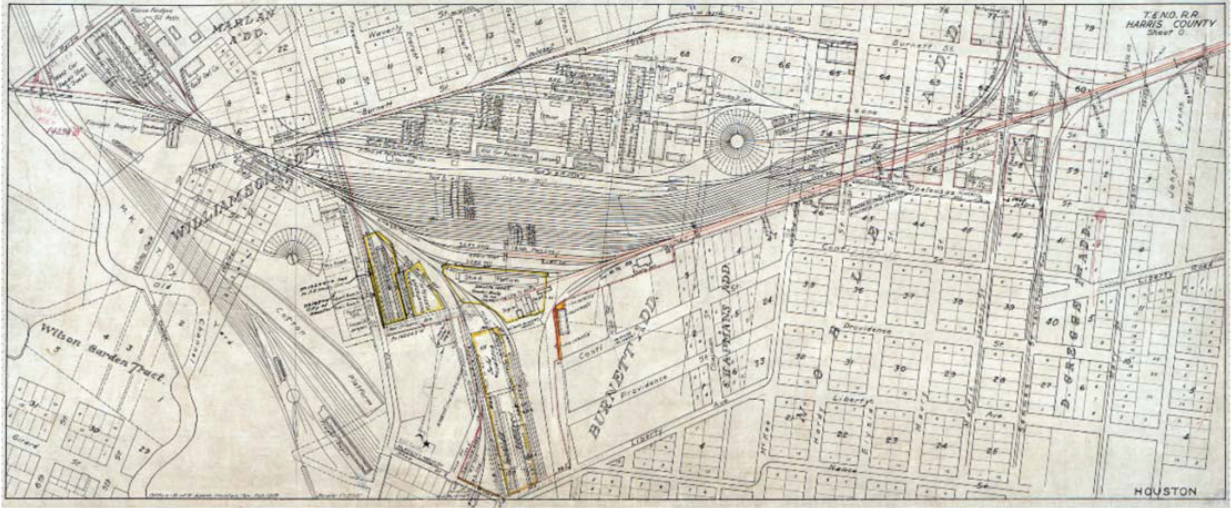


Drag zoom slider to enlarge image

This map shows the routes of 36 railroads in Texas in 1949.



Take A Closer Look



Zoom in to see details of a major
Houston rail yard, c. 1920

Drag zoom slider to enlarge image



Different Draftsmen, Different Styles



Different Draftsmen, Different Styles

No two maps are exactly alike

The maps in this collection were created by different draftsmen for different railroads over a period of several decades. Thus, the styles used to create each map vary from set to set, and sometimes even from map to map within a set. They run the gamut from complex to simple, from artful to workmanlike, but they



Style 1



Style 2



Style 3



Style 4



Style 5



Style 6

Want To Learn More?



Want to Learn More?

More ways to explore this collection



Visit our Flickr site to see the entire collection

This exhibit was created to give you a more in-depth look at the items in this collection. But your experience with these maps doesn't have to stop here!

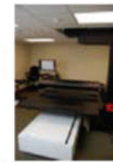
Follow the link on this page to visit our Flickr collection and view all 326 digitized maps. You can search by keywords, or browse from 31 different sets.

Or discover why digitizing these maps helps preserve the information contained on them for future generations.

Plus, you can learn more about the technology used to digitize the maps, or visit the homepage of the Railroad Heritage Museum.



Why were the maps scanned?



Learn how the maps were scanned



Visit the Railroad and Heritage Museum!

Why Were the Maps Scanned?



Why Were the Maps Scanned?

Preserving information, providing access



A damaged map from the collection

The maps in this collection date back as far as the late 1800s, and years of use by railroad staff and researchers alike have taken their toll on them. Drafted in ink on a treated linen paper, these maps were intended for heavy use, but over time the coating has begun to wear thin, and damage spots like those seen on the map at left have eaten holes in some of the objects.

Digitizing these maps allows the archivist at the RHM to place the originals in permanent storage, where they will no longer be available to viewers except under very specific circumstances.

In addition to preserving



See all 326 maps online at
[Flickr.com](https://www.flickr.com/photos/rhm/)

How Were the Maps Scanned?



How Were the Maps Scanned?



Riley Digitization
Center



The Cruse scanner



The Cruse scanner's
light rig

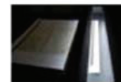
The maps were scanned by Eric Ames, a graduate student in Baylor University's Museum Studies program and an information specialist with the Riley Digitization Center. Using a Cruse CS 285-ST large format scanner, Ames was able to scan 6 maps in one pass. The high resolution images were then preserved on several



The Cruse in action



Riley Digitization Center at Baylor University



Scanning an
architectural drawing

BIBLIOGRAPHY

Cyber Drive Illinois, Illinois State Library. Springfield, IL. Accessed February 8, 2010. Available from http://www.cyberdriveillinois.com/departments/library/who_we_are/bestpractices.html.

Davidson, Lloyd A. et al. "Digital Object Identifiers and Their Role in Implementation of Electronic Publishing", IEEE 1998. Accessed February 18, 2010. Available from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00730709>

"First impressions count for web," BBC News Web site. January 16, 2006. Accessed January 30, 2010. Available from <http://news.bbc.co.uk/2/hi/technology/4616700.stm>.

"Focal Points – Fine Art Reproduction." Epson America. Long Beach, CA. Accessed January 28, 2010. Available from http://www.epson.com/cgi-bin/Store/SuccessStories/ssDetail.jsp?title=ss_mfaBoston&detailType=fineart

Griffiths, Jose-Marie et al. *InterConnections: The IMLS National Study on the Use of Libraries, Museums and the Internet Conclusions Summary*, February 2008.

Guggenheim Museum, Selections from the Collection. New York, NY. Accessed November 20, 2009. Available from (www.guggenheimcollection.org).

Hudson River Valley Heritage, Highland, NY. Accessed February 8, 2010. Available from <http://www.hrvh.org/about/digitalimaging.htm>.

Kelly, Penelope. *Managing digitization projects in a small museum*, Master's project, University of Oregon, 2005. Accessed February 8, 2010. Available from <http://hdl.handle.net/1794/937>

Kenney, Anne R. *Digital to Microfilm Conversion: A Demonstration Project 1994-1996*, Ithaca, NY: Cornell University Library, 2001.

Maher, William. "Copyright: The Archivist and the Law" workshop, sponsored by the Society of American Archivists, held at Texas Christian University, November 5-6, 2009.

Michigan Library Consortium "Copyright and Digitization" Web site. Lansing, MI. Accessed January 20, 2010. Available from http://mlc.lib.mi.us/cms/sitem.cfm/library_tools/copyright_/copyrightdigitize/.

Norris, Chris. "It's Not the Size of the Online Collection ...", *Museum*, January/February 2010. Accessed February 12, 2010. Available from <http://aam-us.org/pubs/mn/onlinecollection.cfm>.

Planning for Library of Congress Collections, *PDF/A-1, PDF for Long-Term Preservation, Use of PDF 1.4*. Washington, DC. Accessed February 10, 2010. Available from <http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml>.

The Texas Heritage Digitization Initiative, Texas State Library and Archives Commission, Austin, TX.
Accessed February 8, 2010. Available at <http://www.tsl.state.tx.us/thdi/>

The UNT Digital Projects Unit. University of North Texas. Denton, TX. Accessed October 8, 2009.
Available from <http://www.library.unt.edu/digitalprojects>.

Van Slyck, Abigail, *Free to All: Carnegie Libraries and American Culture 1890-1920*, Chicago, IL: University of Chicago Press, 1998.

Yan Quan Liu. "Best practices, standards and techniques for digitizing library materials: a snapshot of library digitization practices in the USA", *Online Information Review*, Vo. 28, Issue 5, 2004. Available from
<http://www.emeraldinsight.com/Insight/viewContentItem.do;jsessionid=E6AC16AA8A2F1D034A952C4864323333?contentType=Article&contentId=862266>